

AD-A205 095

OPUS: Optimal Projection for Uncertain Systems

Dennis S. Bernstein, Principal Investigator

DTIC
ELECTE
FEB 15 1989
S & D

$$0 = A_s Q + Q A_s^T + A Q A^T + V_1 + (A - B R_{2s}^{-1} P_s) \hat{Q} (A - B R_{2s}^{-1} P_s)^T - Q_s V_{2s}^{-1} Q_s^T + \tau_1 Q_s V_{2s}^{-1} Q_s^T \tau_1^T$$

$$0 = A_s^T P + P A_s + A^T P A + R_1 + (A - Q_s V_{2s}^{-1} C_s) \hat{P} (A - Q_s V_{2s}^{-1} C_s)^T - P_s R_{2s}^{-1} P_s + \tau_1^T P_s R_{2s}^{-1} P_s \tau_1$$

$$0 = (A_s - B_s R_{2s}^{-1} P_s) \hat{Q} + \hat{Q} (A_s - B_s R_{2s}^{-1} P_s)^T + Q_s V_{2s}^{-1} Q_s^T - \tau_1 Q_s V_{2s}^{-1} Q_s^T \tau_1^T$$

$$0 = (A_s - Q_s V_{2s}^{-1} C_s) \hat{P} + \hat{P} (A_s - Q_s V_{2s}^{-1} C_s) + P_s R_{2s}^{-1} P_s - \tau_1^T P_s R_{2s}^{-1} P_s \tau_1$$

DISTRIBUTION STATEMENT A

Approved for public release
Distribution Unlimited

$$\tau = \sum_{i=1}^{n_i} \Pi_i [\hat{Q} \hat{P}]$$



UNCLASSIFIED

ADA205095

SECURITY CLASSIFICATION OF THIS PAGE

REPORT DOCUMENTATION PAGE

| | | | |
|--|--|---|----------------|
| 1a. REPORT SECURITY CLASSIFICATION | | 1b. RESTRICTIVE MARKINGS | |
| 2a. SECURITY CLASSIFICATION AUTHORITY | | 3. DISTRIBUTION/AVAILABILITY OF REPORT | |
| 2b. DECLASSIFICATION/DOWNGRADING SCHEDULE | | Approved for public release; distribution unlimited. | |
| 4. PERFORMING ORGANIZATION REPORT NUMBER(S) | | 5. MONITORING ORGANIZATION REPORT NUMBER(S) | |
| | | AFOSR-TR-89-0066 | |
| 6a. NAME OF PERFORMING ORGANIZATION | 6b. OFFICE SYMBOL (If applicable) | 7a. NAME OF MONITORING ORGANIZATION | |
| Harris Corporation | | AFOSR/NM | |
| 6c. ADDRESS (City, State and ZIP Code) | | 7b. ADDRESS (City, State and ZIP Code) | |
| Melbourne, FL 32902 | | AFOSR/NM Bldg. 410 Bolling AFB DC 20332-6448 | |
| 8a. NAME OF FUNDING/SPONSORING ORGANIZATION | 8b. OFFICE SYMBOL (If applicable) | 9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER | |
| AFOSR | NM | AFOSR-86-0002 | |
| 8c. ADDRESS (City, State and ZIP Code) | | 10. SOURCE OF FUNDING NOS. | |
| AFOSR/NM BLDG. 410 Bolling AFB DC 20332-6448 | | PROGRAM ELEMENT NO. PROJECT NO. TASK NO. WORK UNIT NO. | |
| 11. TITLE (Include Security Classification) OPUS: Optimal Projection for Uncertain Systems | | 61102F 2304 A1 | |
| 12. PERSONAL AUTHOR(S) Dennis S. Bernstein | | | |
| 13a. TYPE OF REPORT | 13b. TIME COVERED | 14. DATE OF REPORT (Yr., Mo., Day) | 15. PAGE COUNT |
| Final Report | FROM 11/30/88 TO 12/20/88 | OCT 88 | |
| 16. SUPPLEMENTARY NOTATION | | | |
| 17. COSATI CODES | | 18. SUBJECT TERMS (Continue on reverse if necessary and identify by block number) | |
| FIELD | GROUP | SUB. GR. | |
| | | | |
| 19. ABSTRACT (Continue on reverse if necessary and identify by block number) | | | |
| <p>This is the final report for the research project entitled OPUS: Optimal Projection for Uncertain Systems. OPUS is a unified approach to control-system design and analysis for high performance, multivariable applications such as large flexible space structures. In particular, OPUS yields low-order, robust controllers which meet both time- and frequency-domain objectives. The present report is divided into three main research areas:</p> <ol style="list-style-type: none">1) Fixed-Structure Design2) Robust Analysis and Design3) Further Extensions <p>Major accomplishments of the research program include:</p> | | | |
| 20. DISTRIBUTION/AVAILABILITY OF ABSTRACT | | 21. ABSTRACT SECURITY CLASSIFICATION | |
| UNCLASSIFIED/UNLIMITED <input type="checkbox"/> SAME AS RPT. <input type="checkbox"/> DTIC USERS <input type="checkbox"/> | | Unclass | |
| 22a. NAME OF RESPONSIBLE INDIVIDUAL | 22b. TELEPHONE NUMBER (Include Area Code) | 22c. OFFICE SYMBOL | |
| Maj. James M. Crowley | 767-5027 | UNCLASSIFIED | |

(Block #19)

- 1) A unified approach to reduced-order, robust modeling, estimation, and control including singular problems and decentralized architectures
- 2) A computationally tractable approach to designing low-order, finite-dimensional controllers for distributed parameter systems
- 3) A thorough development of quadratic Lyapunov bounds for robust stability and performance analysis
- 4) Complete unification of L_2 (time-domain) and H_∞ (frequency-domain) design criteria for full- and reduced-order modeling, estimation, and control

The report includes reproductions of 41 research papers.

Copy _____ of _____, 5521

Final Report

OPUS: Optimal Projection for Uncertain Systems

For
Air Force Office of Scientific Research (AFSC)
Building 410
Bolling Air Force Base
Washington, DC 20332

Attention:
Major James Crowley

October 1988

This data shall not be disclosed outside the Government or be duplicated, used, or disclosed in whole or in part for any purpose other than to evaluate the proposal, provided that if a contract is awarded to this Offeror as a result of or in connection with the submission of such data, the Government shall have the right to duplicate, use, or disclose this data to the extent provided in the contract. This restriction does not limit the right of the Government to use information contained in such data if it is obtained from another source.

Abstract

This is the final report for the research project entitled OPUS: Optimal Projection for Uncertain Systems. OPUS is a unified approach to control-system design and analysis for high-performance, multivariable applications such as large flexible space structures. In particular, OPUS yields low-order, robust controllers which meet both time- and frequency-domain objectives. The present report is divided into three main research areas:

- 1) Fixed-Structure Design
- 2) Robust Analysis and Design
- 3) Further Extensions

Major accomplishments of the research program include:

- 1) A unified approach to reduced-order, robust modeling, estimation, and control including singular problems and decentralized architectures
- 2) A computationally tractable approach to designing low-order, finite-dimensional controllers for distributed parameter systems
- 3) A thorough development of quadratic Lyapunov bounds for robust stability and performance analysis
- 4) Complete unification of L_2 (time-domain) and H_∞ (frequency-domain) design criteria for full- and reduced-order modeling, estimation, and control. (16)

The report includes reproductions of 41 research papers.



| | |
|--------------------|-------------------------------------|
| Accession For: | |
| NTIS CRA&I | <input checked="" type="checkbox"/> |
| DTIC TAB | <input type="checkbox"/> |
| Unannounced | <input type="checkbox"/> |
| Justification | |
| By | |
| Distribution / | |
| Availability Codes | |
| Dist | Avail and/or Special |
| A-1 | |

Table of Contents

| <u>Section</u> | <u>Title</u> | <u>Page</u> |
|----------------|--|-------------|
| 1.0 | INTRODUCTION. | 1-1 |
| 1.1 | Overview. | 1-1 |
| 1.2 | Status of Computational Results | 1-4 |
| 1.3 | Long-Range Goals of the Project | 1-6 |
| 1.4 | Plan of This Report | 1-6 |
| 2.0 | FIXED-STRUCTURE DESIGN. | 2-1 |
| 2.1 | Motivation. | 2-1 |
| 2.2 | The Three Basic Problems. | 2-1 |
| 2.3 | Finite-Dimensional Control of Distributed Parameter Systems | 2-2 |
| 2.4 | Decentralized Control | 2-5 |
| 2.5 | Singular Control. | 2-7 |
| 3.0 | ROBUST ANALYSIS AND DESIGN. | 3-1 |
| 3.1 | Motivation. | 3-1 |
| 3.2 | Stochastic Modeling | 3-1 |
| 3.3 | Robust Analysis | 3-3 |
| 3.4 | Robust Synthesis. | 3-6 |
| 3.5 | H_∞ Theory | 3-7 |
| 4.0 | FURTHER EXTENSIONS. | 4-1 |
| 4.1 | Motivation. | 4-1 |
| 4.2 | Tracking. | 4-1 |
| 4.3 | Sampled-Data Control. | 4-2 |
| 5.0 | OPEN PROBLEMS | 5-1 |
| 5.1 | Motivation. | 5-1 |
| 5.2 | Fixed Structure Design. | 5-1 |
| 5.3 | Robust Analysis and Design. | 5-2 |
| 5.4 | Tracking and Sampled-Data Control | 5-3 |
| 6.0 | COMPREHENSIVE REFERENCE LIST. | 6-1 |
| 7.0 | PROGRAM PERSONNEL | 7-1 |
| 7.1 | Dr. Dennis S. Bernstein | 7-1 |
| 7.2 | Professor Wassim M. Haddad. | 7-1 |
| 7.3 | Acknowledgements. | 7-1 |

List of Appendices

| <u>Appendix</u> | <u>Title</u> | <u>References</u> |
|-----------------|--|-----------------------|
| A | OPUS Review Paper | 88 |
| B | Fixed Structure Design. | 32,29,24 |
| C | Distributed Parameter Systems | 37,122 |
| D | Decentralized Control | 76,121 |
| E | Singular Control. | 78,79,130 |
| F | Stochastic Modeling | 35,36,104,39,77 |
| G | Robust Analysis | 115,123,75 |
| H | Robust Synthesis: Linear Bound | 95,119,89 |
| I | Robust Synthesis: Quadratic Bound. | 101,83,94,113 |
| J | H_{∞} Theory | 114,116,117,115 |
| K | Tracking Control. | 67,86,103,125 |
| L | Discrete-Time Theory. | 41,44,54,69,92,87,128 |

SECTION 1.0
INTRODUCTION

1.0 INTRODUCTION

1.1 Overview

Over the past 10-15 years controls researchers have come to the realization that classical controls analysis and design techniques are inadequate in the face of modern large scale, high-performance applications. In particular, the principal motivation for OPUS is the problem of vibration suppression in large lightweight flexible space structures characterized by high-dimensional, highly uncertain models. In addition, stringent performance specifications in the face of high disturbance levels place severe demands on existing control-design techniques. Specifically, performance tradeoffs involving sensors, processors, actuators, and identification accuracy must be cut as tightly as possible to minimize hardware and testing costs. For feasibility and cost effectiveness, system design must also be performed efficiently with respect to human and computer resources.

The goal of this project has been to develop a mathematically rigorous control-design methodology which directly addresses these technology issues. In particular, optimal projection theory addresses the need for low-order, high-performance controllers which can be implemented on-board for real-time operation. Low-order controllers are necessitated by cost, weight, and reliability constraints associated with space-qualified processors. Furthermore, OPUS incorporates a fundamental theory of robust controller synthesis to account for unavoidable modeling uncertainties arising for reasons such as material and manufacturing variations, thermal and aging effects, as well as limits to identification accuracy. The principal contribution of OPUS is thus a unified theory which simultaneously accounts for both real-time processor constraints and modeling uncertainty. A high level overview of OPUS is given in [88] (Appendix A).

During the course of this project OPUS has, in addition, been extended to a large class of problems in systems and control theory. The current scope of the theory includes (see Figure 1-1):

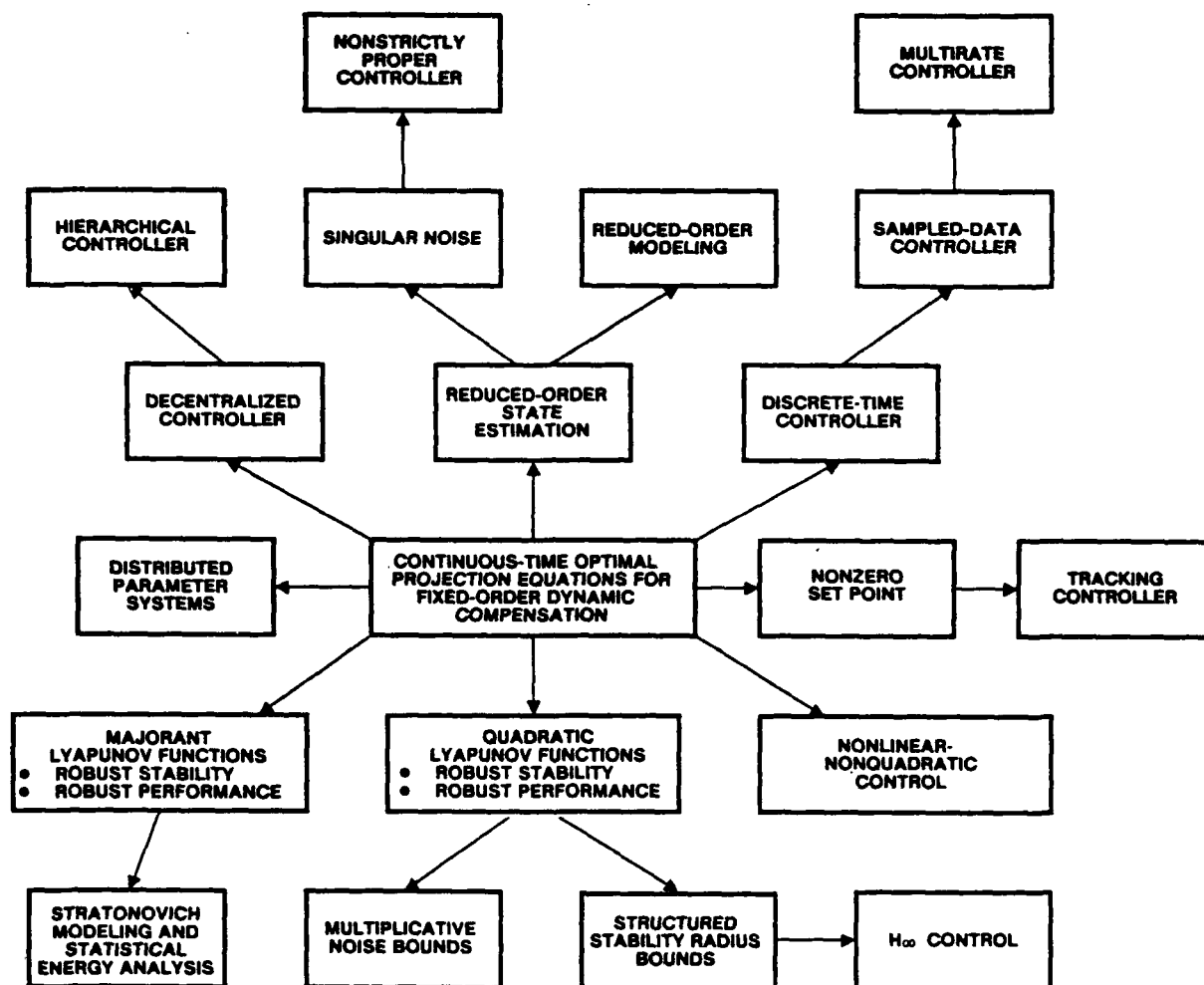


Figure 1-1. Scope of OPUS

1. A unified treatment of reduced-order modeling, estimation, and control (Appendix B);
2. Robust estimation and control via quadratic Lyapunov functions including robust performance (Appendices G,H,I);
3. A unified approach to L_2 and H_∞ control including parametric robustness (Appendix J);
4. Decentralized, nonstrictly proper, and sampled-data control (Appendices D,E,L).

Of particular interest is the recent extension to H_∞ control. As shown in [117] (Appendix J), we have developed a method for directly imbedding H_∞ design constraints within OPUS theory and thus, in particular, within LQG. These results are given by a system of modified Riccati equations which directly generalize LQG theory and which have the potential for significant computational savings compared to existing H_∞ synthesis methods.

The underlying philosophy of OPUS is to capture as many design constraints as possible within a single system of design equations. This is demonstrated in [117] by the unification of time- and frequency-domain criteria addressed by the L_2/H_∞ design equations. An additional example is provided by the results obtained in [119,94] (Appendices H and I) for robust stability and performance via fixed-order compensation in the presence of real-valued structured parameter uncertainty. In these algebraic design equations the projection matrix automatically enforces a constraint on controller order, while additional terms guarantee both robust stability and performance. Note that for full-order controllers in the absence of uncertainty, these four coupled equations reduce to the standard pair of separated Riccati equations of LQG theory. Versions of these equations have been developed for each of the problems shown in Figure 1-1. These results are discussed in more technical detail in the following sections.

The justification for this line of research is based upon several considerations. First, and most obvious, is the fact that our results show that numerous design constraints can be captured simultaneously within a constructive theory which directly generalizes LQG theory. Such an approach provides the capability for simultaneously performing multiple design tradeoffs

for multivariable systems with respect to competing constraints such as sensor noise, control authority, controller order, robustness, disturbance attenuation, mean-square error, sample rate, degree of decentralization, etc. Next we stress that rather than being ad hoc constructions, these design equations follow directly from the optimality of well-defined performance objectives. Thus, these results are useful in assessing the suboptimality of alternative methods. For example, as shown in [20] several suboptimal approaches to reduced-order control design can be viewed as approximations to the optimal projection equations.

1.2 Status of Computational Results

Overall, OPUS can be viewed as a theory for characterizing solutions to constrained control-design problems. Transforming OPUS into a practical design methodology requires the development of effective computational algorithms. Such development has been carried out in related work by S. Richter at Harris Corporation. Using homotopic continuation methods, Richter has developed efficient algorithms which fully account for the structure of these modified Riccati equations and their coupling terms. Homotopy algorithms, as reviewed in

S. Richter and R. DeCarlo, "Continuation Methods: Theory and Applications," IEEE Trans. Autom. Contr., Vol. 28, pp. 660-665, 1983.

offer several advantages over both gradient-based and Newton-type methods. For example, homotopy methods have a strong theoretical foundation based upon differential topology, in particular, topological degree theory, while in practice these methods effectively address the key issues of startup, convergence, and global optimality. Homotopy algorithms have also reached a high degree of maturity and availability with the advent of HOMPACK described in

L. T. Watson, S. C. Billups, and A. P. Morgan, "A Suite of Codes for Globally Convergent Homotopy Algorithms," ACM Trans. Math. Software, Vol. 13, pp. 281-310, 1987.

The continuation algorithm developed for the optimal projection equations essentially follows a smooth path connecting an easily solvable version of the equations with the final, desired form. The algorithm utilizes the tensor derivatives of the terms in the optimal projection equations to integrate along the solution paths. To demonstrate the algorithm, an 8th-order, nonminimum phase example originally due to

R. H. Cannon, Jr., and D. E. Rosenthal, "Experiments in Control of Flexible Structures with Noncollocated Sensors and Actuators," AIAA J. Guid. Contr. Dyn., Vol. 7, pp. 546-553, 1984.

was considered. This problem was used in

Y. Liu and B. D. O. Anderson, "Controller Reduction Via Stable Factorization and Balancing," Int. J. Contr., Vol. 44, pp. 507-531, 1986.

to compare several reduced-order control-design methods. The comparisons performed by Liu and Anderson highlight the suboptimal nature of these methods. Specifically, several methods failed to yield stabilizing controllers for 10% of the cases while others failed for as many as 60%. In contrast, as reported in [68,102], the optimal projection approach yielded stabilizing controllers for all cases considered. While the methods compared by Liu and Anderson were most prone to failure at high authority levels, the optimal projection results were within 20% of the LQG performance at 10^2 - 10^3 higher authority levels. In addition, using topological degree theory, an upper bound has been obtained on the number of solutions of the design equations. Letting n = plant dimension, n_u = dimension of the unstable plant subspace, n_c = compensator order, l = number of measurements, and m = number of controls, the number of solutions for the case $n_c \geq n_u$ is not greater than

$$\begin{pmatrix} \min(n,m,\ell) - n_u \\ n_c - n_u \\ 1 \end{pmatrix}, \quad n_c \leq \min(n,m,\ell),$$

, otherwise.

Hence, for the case in which the controller order is greater than the number of inputs or outputs (so that the controller is not ill-conditioned), the equations possess at most one solution corresponding to the global minimum. Furthermore, since in many practical cases of interest this bound is small, it suffices to compute each such solution to determine the global optimum. These results along with suitable extensions to related problems have been used widely throughout this project. For example, recent results on fixed-order control of distributed parameter systems described in Section 2.3 were obtained using the homotopy algorithm.

1.3 Long-Range Goals of the Project

The long-range (5-10 year) goal of this project is the development of a truly effective computer-aided design methodology for multivariable control design. Numerical solution of the design equations would form the basis for such a design tool. This methodology would be appropriate in an engineering environment since the user need not be familiar with the mathematics of the design equations being solved. We envision a methodology similar to finite element modeling used routinely by structural analysts. An OPUS design package would go far beyond currently available packages whose multivariable design capabilities are based largely upon LQG theory.

1.4 Plan of This Report

Since this is the final report for this project our goal is to accomplish the following objectives:

- 1) Review the evolution and maturation of the research plan throughout the project;
- 2) Highlight the principal research accomplishments; and

- 3) Summarize open problems and point out future research directions.

Detailed technical discussion of results obtained will not appear in the main body of the report. Rather, the appendices contain a fairly complete (and lengthy) collection of the principal research results. We note that the ordering of the appendices is not chronological but instead reflects the most logical order according to subject matter.

SECTION 2.0
FIXED-STRUCTURE DESIGN

2.0 FIXED-STRUCTURE DESIGN

2.1 Motivation

While achieving the system design specifications (stability, performance, etc.) the control-design process must not lose sight of restrictions which arise in controller implementation. Indeed, control-design methods which focus primarily on performance specifications often pay a serious price by producing controllers which are difficult, if not impossible, to implement in practice. Hence our approach rests upon the notion of fixed-structure design. That is, we seek to meet design specifications within a framework which constrains the class of implementable designs. In this way the burden of hardware implementation (sensors, processors, and actuators) can be minimized to the greatest possible extent.

2.2 The Three Basic Problems

The most fundamental restriction arising in fixed-structure design is that of the order, or dimension, of the controller. In addressing this problem we have developed a unified treatment of three basic problems in reduced-order design, namely, modeling, estimation, and control. These three problems form a fundamental hierarchy of design problems in system theory, namely, to determine a system of fixed degree which, for a given system, approximates, estimates, or controls selected plant states. The solutions to these problems, given in [32,29,24] (Appendix B), reveal a surprising degree of common structure. Specifically, the solutions involve systems of 2, 3, and 4 modified algebraic Riccati and Lyapunov equations coupled by a projection matrix (the "optimal projection"). In addition, the estimation and control results provide transparent generalizations of steady-state Kalman filter and LQG theory.

Although the structure of these equations is aesthetically appealing by itself, the principal benefit for practical purposes is computational. That is, by exploiting the structure of these equations it is possible to significantly reduce the computational burden inherent in commonly used

gradient search techniques. This point has been amply demonstrated in [63,68] as discussed in Section 1.

2.3 Finite-Dimensional Control of Distributed Parameter Systems

The problem of controller order becomes exacerbated when the plant is infinite dimensional since infinite-dimensional controllers cannot be implemented precisely, while finite-dimensional plant approximations may be of arbitrarily high order. To address this problem the fixed-structure control-design results of [24] were generalized in [37] (Appendix C) to the case in which the plant is infinite dimensional. The resulting design equations now comprise a system of four operator equations coupled by a finite-rank nonselfadjoint projection operator. In spite of the infinite dimensionality of the plant, the design equations directly characterize fixed-order, finite-dimensional dynamic compensator gains (Figure 2-1). Corresponding results for fixed-order finite-dimensional modeling and fixed-order finite-dimensional state estimation can also be obtained in an analogous manner.

Application of the operator-theoretic results of [37], however, requires finite-dimensional approximation of the design equations. In practice one could solve the design equations for a sequence of plant approximations of increasingly high order while keeping the controller order fixed. The limiting controller would then serve as a nearly optimal fixed-order finite-dimensional controller for the original distributed parameter system (Figure 2-2). This was investigated numerically in [122] in a collaborative project with Professor I. G. Rosen. In [122] two alternative approaches were considered for obtaining finite-dimensional controllers for infinite-dimensional systems. The first approach, which has been widely studied, involves computing a sequence of full-order LQG controllers for a sequence of high-order plant approximations, while the second approach assumes a fixed order for the dynamic controller. To demonstrate these methods, two examples were considered, namely, a one-dimensional parabolic (heat/diffusion) system and a hereditary (delay) system. For each example a sequence of spline-based, Ritz-Galerkin finite element approximations was derived for use in the control-design procedure. LQG theory and the optimal projection approach were then used to obtain full- and first-

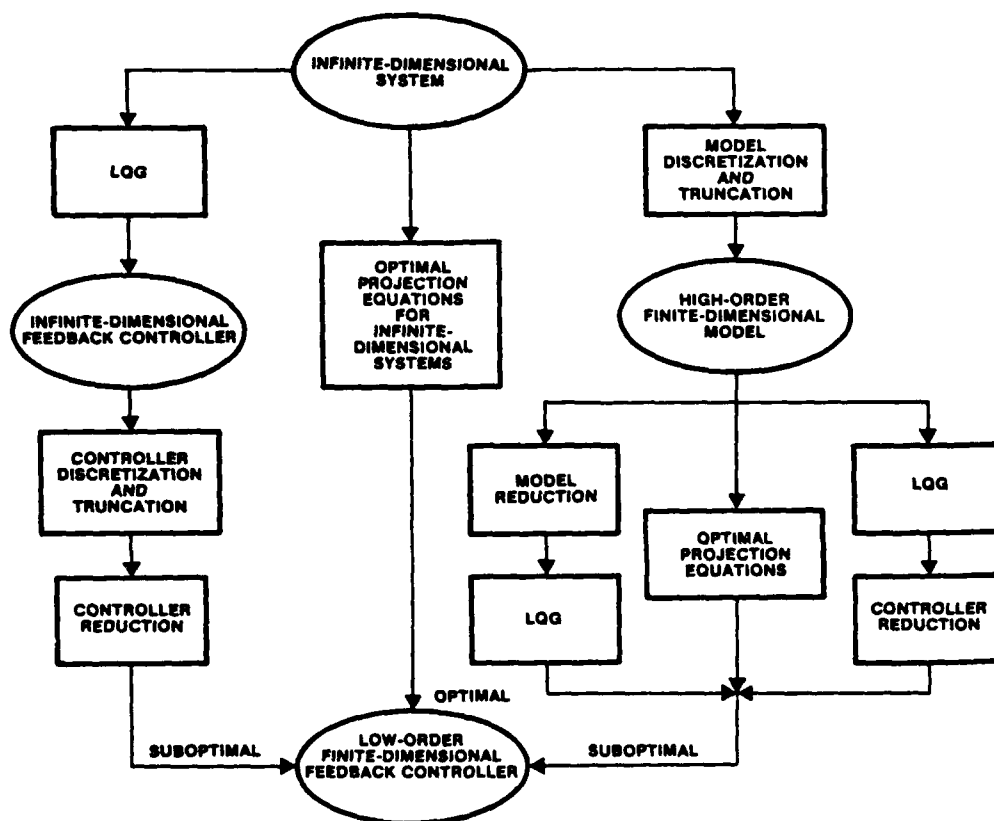


Figure 2-1. The Optimal Projection Equations For Finite-Dimensional Fixed-Order Dynamic Compensation of Infinite-Dimensional Systems Provide a Direct Path to Optimal Physically Realizable Controllers for Distributed Parameter Systems

CONVERGENCE OF SUBOPTIMAL REDUCED-ORDER COMPENSATORS

IDEA: DESIGN A SEQUENCE OF REDUCED-ORDER COMPENSATORS WHILE INCREASING THE ORDER OF THE APPROXIMATE MODEL AND KEEPING THE ORDER OF THE COMPENSATOR FIXED

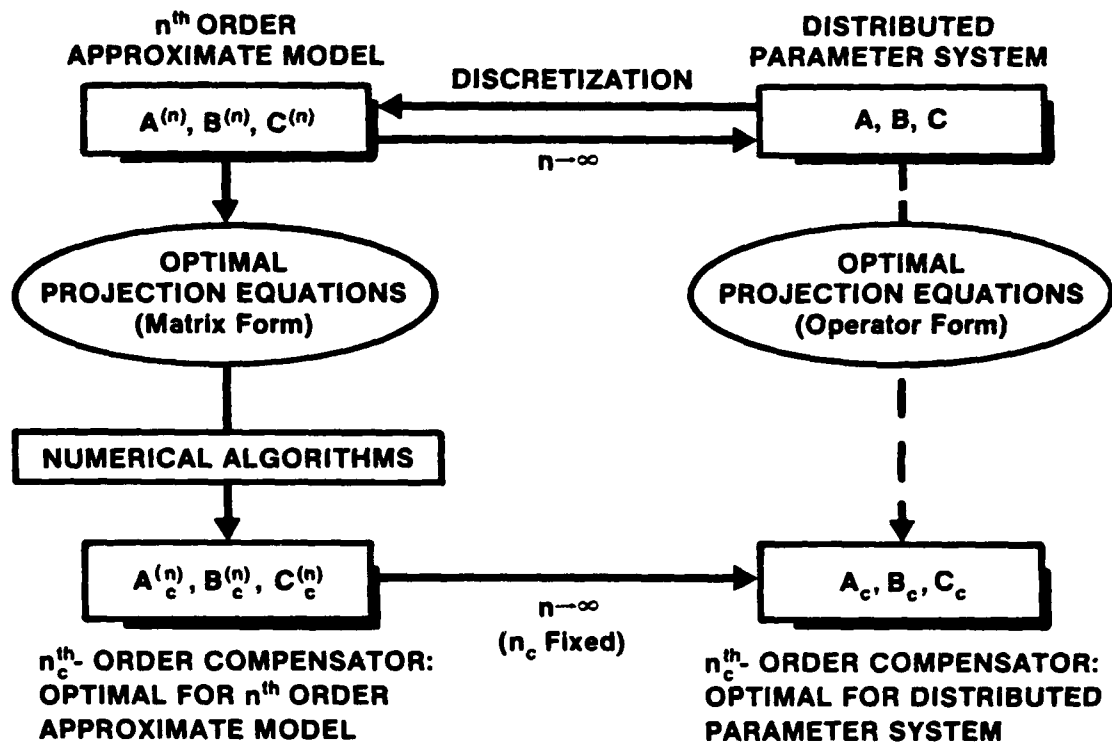


Figure 2-2. Numerical Solution of the Optimal Projection Equations for Fixed-Order Dynamic Compensation Provides a Path to the Optimal Fixed-Order Controller for an Infinite-Dimensional System

order controllers for each example with plant approximations up to 32nd order. For the parabolic system the performance degradation of the first-order controllers was only 2% compared to the full-order controllers (Figure 2-3), while for the hereditary system the degradation was less than 10%. The difference in implementation requirements for a first-order versus a 32nd-order controller is, of course, considerable.

2.4 Decentralized Control

In addition to incorporating constraints on the order of the feedback compensator, the fixed-structure approach allows additional constraints on the complexity of the feedback law. In particular, the results of [24] assumed a centralized structure for the dynamic compensator. In many applications, however, a decentralized controller architecture permits a simplified feedback communication structure and allows increased parallelism in the control law execution.

The fixed-structure approach is ideally suited to the decentralized design problem. For each fixed decentralized architecture, the design procedure can be performed to assess the ability to meet specifications for the given configuration. If specifications cannot be met, then the feedback architecture can be modified to improve performance, robustness, etc.

For the case of dynamic compensation, it was shown in [76] that the optimal projection technique provides a direct means for characterizing decentralized controllers. The key step is the realization that each subcontroller in the decentralized configuration must be an optimal centralized controller when viewed as a controller for the plant and remaining subcontrollers. This observation immediately suggests a sequential design algorithm in which individual subcontrollers are alternately refined until convergence is achieved. Because the method is based upon optimization principles, performance improvement is guaranteed at each step. This technique was demonstrated numerically in [76] (Appendix D) where a two-channel decentralized controller, fourth-order in each channel, was designed for a pair of interconnected simply supported beams. The algorithm demonstrated

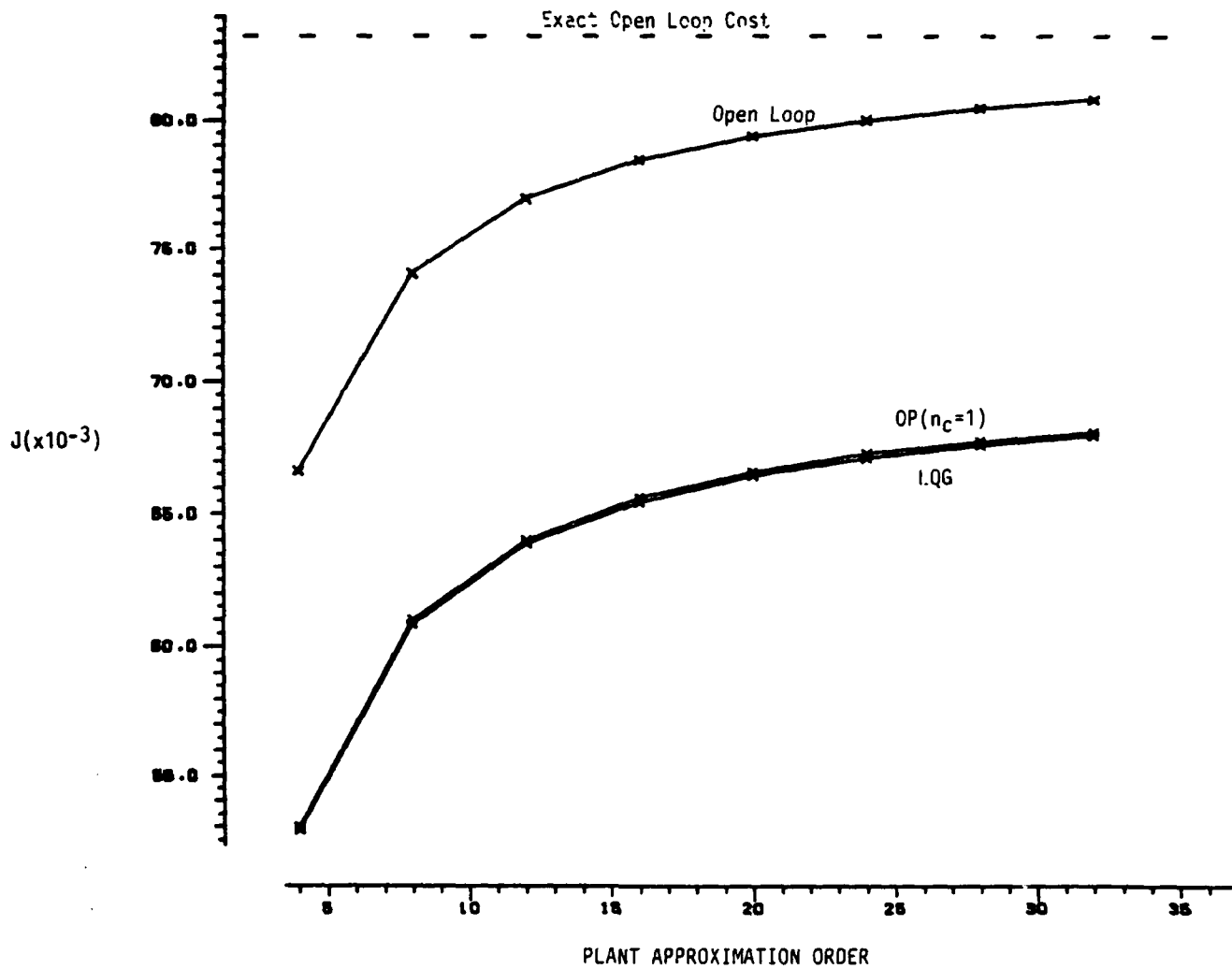


Figure 2-3. The LQG Closed-Loop Cost via Full-Order Controllers is Compared to First-Order Optimal Projection Designs for a Parabolic System

convergence to a decentralized controller whose performance was within 10% of the fully centralized controller.

For the case in which each subcontroller is a static (proportional) feedback law, it is possible to simultaneously characterize the optimal gains in each control channel without requiring a sequential approach. A thorough treatment of this case, including robust stability and performance, is given in [121] (Appendix D).

2.5 Singular Control

An important generalization of the results of [24] involves the case in which the controller includes a static feedthrough component. One technical issue which arises in the problem formulation is that the L_2 norm of a control signal corrupted by white noise (as a result of measurement feedthrough) is infinite. Hence the measurement feedthrough problem is only well-posed when either the measurement noise intensity or the control weighting matrix is singular. As is well known from the singular control literature, however, singular problem data often lead to complex behavior including impulsive controls and singular arcs. The imposition of a smooth controller structure via the fixed-structure approach thus precludes such complex behavior.

The fixed-order state estimation and dynamic compensation results of [29,24] were partially extended to the singular case in [78,79]. Even in the full-order case the singular control results are novel since standard LQG theory yields only strictly proper controllers. The results of [78,79] were incomplete, however, since the gains associated with certain estimation and feedback paths were not given explicitly. For the singular estimation problem this defect was remedied in

Y. Halevi, "The Optimal Reduced-Order Estimator for Systems with Singular Measurement Noise," IEEE Trans. Autom. Contr., Vol. 34, 1989.

where all feedback gains were explicitly characterized. In addition, this

solution was shown to agree completely with results obtained using standard limiting methods. For the corresponding dynamic-compensation problem the complete singular solution has been derived in joint research with Professor Y. Halevi and will be reported in [130,138,139] (Appendix E).

SECTION 3.0
ROBUST ANALYSIS AND DESIGN

3.0 ROBUST ANALYSIS AND DESIGN

3.1 Motivation

The purpose of feedback control is to achieve desirable performance in the face of uncertainty. Although identification can reduce uncertainty to some extent, it is often impractical and residual modeling discrepancies always remain. For example, modeling uncertainty in flexible structures may arise in the mass, damping, and stiffness operators. Controllers must therefore be robust to achieve desired disturbance rejection in the presence of such modeling uncertainty.

3.2 Stochastic Modeling

Our approach to robust control was originally inspired by stochastic parameter modeling within a linear-quadratic optimization framework. In a series of early papers [1-16], D. C. Hyland explored the ramifications of a multiplicative white noise model as a consequence of the minimum information modeling technique based upon the Maximum Entropy Principle of Jaynes. The intent was not to view the white noise process as a literal model of parameter uncertainty, however, but rather to construct a tractable design model which captures the effects of parameter uncertainty upon system performance.

An interesting feature of the Maximum Entropy modeling approach was that the multiplicative white noise model was not to be rigorously interpreted as an Ito differential model, but rather in terms of the Stratonovich formulation. Recasting the Stratonovich model in terms of Ito differentials then led to additional "correction" terms. It is precisely these terms which were shown to play a crucial role in capturing the effects of parameter uncertainty. Such effects include decorrelation, i.e., the decrease in cross-correlation of system states due to parameter uncertainty, as well as equilibration, i.e., the tendency of state variances to equalize in the presence of high levels of uncertainty thus rendering different states indistinguishable. These effects of parameter uncertainty are fundamental features of high-order, lightly damped modal systems. An interesting treatment of these ideas for structural and

acoustic analysis can be found in

R. H. Lyon, Statistical Energy Analysis of Dynamical Systems: Theory and Application, MIT Press, Cambridge, MA, 1975.

For feedback design within fixed-structure design theory, the Stratonovich model produces controllers possessing intuitively appealing features. Specifically, such control laws exhibit high-authority control in the low-frequency, well-modeled portion of the structure along with low-authority, rate dissipative action in the high-frequency region [35] (Appendix F). The ability to merge and unify these control regimes is a unique and significant contribution of the Maximum Entropy approach.

As a control-design methodology, however, it remained to validate the approach as a rigorous robust design technique. Optimal controllers designed in the presence of white noise disturbances, it was reasoned, are automatically desensitized to actual deterministic plant parameter variations. This idea was confirmed empirically by numerical studies in [36,39] which showed an efficient design tradeoff between performance and robustness in the presence of structured real-valued parameter variations. Further robustness studies confirming these results were carried out in

A. Gruzen, "Robust Reduced Order Control of Flexible Structures," C. S. Draper Laboratory Report #CSDL-T900, April 1986.

A. Gruzen and W. E. Vander Velde, "Robust Reduced-Order Control of Flexible Structures Using the Optimal Projection/Maximum Entropy Design Methodology," AIAA Guid. Nav. Contr. Conf., Williamsburg, VA, August 1986.

M. Cheung and S. Yurkovich, "On the Robustness of MEOP Design versus Asymptotic LQG Synthesis," IEEE Trans. Autom. Contr., Vol. 33, 1988.

In spite of these results, it was clear that issues concerning stochastic modeling, such as stochastic stability and the physical interpretation of the model, tended to obscure the effectiveness of the technique for robust control. Thus a crucial step in the evolution of our approach was the ability

to show in [77] (Appendix F) that such controllers are guaranteed to be robust for all cases in which the design equations are solvable. In particular, it was shown that a second-moment stochastic stability condition in the presence of a time-exponential cost weighting induces a Lyapunov function which guarantees deterministic robust stability over a prescribed range of parameter variations. This result thus provided the bridge to cross over from the world of stochastic modeling (a statistical theory) to deterministic robustness theory (a theory of worst-case bounds).

3.3 Robust Analysis

For a given controller, it is often necessary to assess the stability and worst-case performance of the closed-loop system as parameters vary within a specified range of uncertainty. This is a problem of robust analysis, whose consideration precedes the more complex problem of robust controller synthesis.

Our principal mathematical technique in robustness analysis is Lyapunov stability theory. Here the idea is to determine a Lyapunov function which guarantees robust stability over a range of uncertain parameters. For linear systems we employ the quadratic Lyapunov function

$$V(x) = x^T P x \quad (1)$$

or, equivalently, the Lyapunov equation

$$0 = A^T P + P A + R \quad (2)$$

for the linear system

$$\dot{x} = A x + w. \quad (3)$$

The dual equation

$$0 = A Q + Q A^T + V \quad (4)$$

is also useful for robust performance analysis since V can be interpreted as the intensity of the additive white noise signal w . In robust analysis one typically replaces (4) by

$$0 = A\bar{Q} + \bar{Q}A^T + \Omega + V, \quad (5)$$

where Ω is an additional nonnegative-definite matrix. Now robust stability of the perturbed system

$$\dot{x} = (A+\Delta A)x + w \quad (6)$$

is assured so long as

$$\Delta A\bar{Q} + \bar{Q}\Delta A^T < \Omega. \quad (7)$$

This can be seen by rewriting (5) as

$$0 = (A+\Delta A)\bar{Q} + \bar{Q}(A+\Delta A)^T + [\Omega - (\Delta A\bar{Q} + \bar{Q}\Delta A^T)] + V. \quad (8)$$

Furthermore, it is also possible to guarantee robust performance since the solution $Q_{\Delta A}$ of

$$0 = (A+\Delta A)Q_{\Delta A} + Q_{\Delta A}(A+\Delta A)^T + V \quad (9)$$

satisfies

$$Q_{\Delta A} \leq \bar{Q}. \quad (10)$$

The above technique, developed in [115] (Appendix G), provides a simple

approach to robust stability and performance.

To develop a more sophisticated approach one can replace (5) by

$$0 = AQ + QA^T + \Omega(Q) + V \quad (11)$$

where $\Omega(\cdot)$ is now a bounding operator which satisfies

$$\Delta AQ + Q\Delta A^T < \Omega(Q) \quad (12)$$

for all variations ΔA in a specified uncertainty set and for all nonnegative-definite matrices Q . This approach now guarantees the bounding a priori via (12) and the problem is to determine whether or not there exists a solution to (11).

The a priori bounding technique shown in (11), (12) has been given a fairly complete treatment in [123] (Appendix G). The goal in [123] was to systematically investigate candidate choices for the function $\Omega(\cdot)$. This investigation also provides a unified setting for particular bounds which have been used in various control-design contexts. For example, for $A = \sigma_1 A_1$, $|\sigma_1| \leq 1$, the absolute value bound

$$\Omega(Q) = |A_1 Q + Q A_1^T|, \quad (13)$$

where $|\cdot|$ replaces each eigenvalue by its absolute value, was proposed in

S. S. L. Chang and T. K. C. Peng, "Adaptive Guaranteed Cost Control of Systems with Uncertain Parameters", IEEE Trans. Autom. Contr., Vol. AC-17, pp. 474-483, 1972.

On the other hand, writing $A_1 = D_1 E_1$, the bound

$$\Omega(Q) = D + Q E Q, \quad (14)$$

where $D = D_1 D_1^T$ and $E = E_1^T E_1$, was studied in

I. R. Petersen and C. V. Hollot, "A Riccati Equation Approach to the Stabilization of Uncertain Systems", Automatica, Vol. 22, pp. 433-448, 1986.

D. Hinrichsen and A. J. Pritchard, "Stability Radius for Structured Perturbations and the Algebraic Riccati Equation", Sys. Contr. Lett., Vol. 8, pp. 105-113, 1986.

Finally, the choice

$$\Omega(Q) = \alpha Q + \alpha^{-1} A_1 Q A_1^T \quad (15)$$

corresponds to the bound arising from a multiplicative white noise model as discussed in [77] and Section 3.2. We call (14) the quadratic bound (since it is quadratic in Q) and (15) the linear bound (since it is linear in Q).

3.4 Robust Synthesis

The principal payoff of our robust stability and performance technique is the ability to incorporate these bounds directly within the fixed structure design methodology. This can be done by bounding the cost over the class of parameter uncertainties prior to determining the feedback gains. The resulting bound is then treated as an auxiliary cost which can then be minimized by suitable feedback gains. The solution to this optimization problem is thus guaranteed to yield robust stability and performance.

To carry out this procedure it is essential that the bound $\Omega(\cdot)$ be differentiable with respect to Q . Furthermore, $\Omega(\cdot)$ will be differentiable with respect to the feedback gains if it is differentiable with respect to A_1 (which involves gains in the control-design setting). These requirements thus suggest the linear bound (15) and the quadratic bound (14) as the prime candidates for robust synthesis.

As discussed previously, the linear bound (15) was originally suggested by a multiplicative white noise model. By incorporating this bound within the design procedure, sufficient conditions for robust estimation and robust control were developed in [95,119] (Appendix H). In addition, a unified treatment of robust, reduced-order modeling, estimation, and control was given in [89] (Appendix H).

The quadratic bound (14) has also been developed extensively within a design context. In [101,83,94] (Appendix I) the problems of reduced-order modeling, estimation, and control were considered via this bound. Finally, both the linear and quadratic bounds were considered simultaneously in [113] (Appendix I).

3.5 H_∞ Theory

The robustness theory discussed in the previous subsections addresses the problem of real-valued structured parameter uncertainty. In many applications, however, uncertainty is present in the form of unstructured perturbations to the plant transfer function. A typical case is the presence of high-frequency, unmodeled dynamics.

A mathematically rigorous approach to this problem involves defining a suitable norm on the space of plant transfer functions to characterize uncertainty in terms of neighborhoods of the nominal plant. The resulting H_∞ theory was pioneered by Zames in

G. Zames, "Feedback and Optimal Sensitivity: Model Reference Transformations, Multiplicative Seminorms, and Approximate Inverses," IEEE Trans. Autom. Contr., Vol. AC-26, pp. 301-320, 1981.

while recent overviews were given in

B. A. Francis and J. C. Doyle, "Linear Control Theory with an H_∞ Optimality Criterion," SIAM J. Contr. Optim., Vol. 25, pp. 815-844, 1987.

B. A. Francis, A Course in H_∞ Control Theory, Springer-Verlag, New York, 1987.

The most fundamental problem of H_∞ control design is the so-called Standard Problem considered by Francis: determine a feedback compensator which minimizes the peak (worst-case) disturbance attenuation of the closed-loop system. By introducing suitable weighting matrices and problem transformations, solutions to the Standard Problem can be used to yield robust controllers for unstructured plant uncertainty.

Current H_∞ synthesis methods, however, possess two principal drawbacks: they are computationally intensive and they often yield excessively high-order controllers. These difficulties have been removed with the advent of new state space solutions to the Standard Problem given in [117] (Appendix J) and

I. R. Petersen, "Disturbance Attenuation and H^∞ Optimization: A Design Method Based on the Algebraic Riccati equation," IEEE Trans. Autom. Contr., Vol. AC-32, pp. 427-429, 1987.

P. P. Khargonekar, I. R. Petersen, and M. A. Rotea, " H^∞ Optimal Control with State Feedback," IEEE Trans. Autom. Contr., Vol. 33, pp. 786-788, 1988.

J. C. Doyle, K. Glover, P. P. Khargonekar, and B. A. Francis, "State-Space Solutions to Standard H_2 and H_∞ Control Problems," Proc. Amer. Contr. Conf., pp. 1691-1696, Atlanta, GA, June 1988.

These papers characterize solutions to the Standard Problem in terms of modified Riccati equations. The computational savings of this approach over earlier methods is considerable, possibly two orders of magnitude. In addition, the dynamic compensators obtained from these Riccati equations are of the same order as the plant model. This approach thus removes the principal drawbacks of earlier H_∞ synthesis methods.

By incorporating the fixed-structure approach we have, in addition, obtained the most general solution thus far available for the Standard Problem. Specifically, in [117] (Appendix J) we consider the minimization of an L_2 performance criterion subject to a constraint on the H_∞ closed-loop performance. This multi-norm problem formulation thus allows the designer to

perform tradeoffs between these competing performance measures. In addition we impose a constraint on the order of the dynamic compensator to obtain optimal low-order feedback controllers which satisfy the H_∞ performance constraint. Utilizing an eighth-order nonminimum phase example given in

R. H. Cannon, Jr., and D. E. Rosenthal, "Experiments in Control of Flexible Structures with Noncollocated Sensors and Actuators," AIAA J. Guid. Contr. Dyn., Vol. 7, pp. 546-553, 1984.

we used these results to obtain 9 dB improvement over the corresponding LQG design (Figure 3-1).

Immediate spinoffs of these results include the problems of model reduction and state estimation. The H_∞ model reduction problem [114] (Appendix J) addresses one of the most fundamental problems of linear system theory, namely, given a linear time-invariant system of degree n , find a linear time-invariant transfer function of degree $n_m < n$ which minimizes the H_∞ distance between the full- and reduced-order systems. Although the Hankel norm model-reduction problem has been widely studied as in

K. Glover, "All Optimal Hankel-Norm Approximations of Linear Multivariable Systems and Their L^∞ -Error Bounds," Int. J. Contr., Vol. 39, pp. 1115-1193, 1984.

the solution to the H_∞ problem had not been given previously.

For state estimation the Kalman filter provides the least squares (L_2) optimal solution. In certain applications, however, it may be desirable to minimize the worst-case frequency content of the error signal. This problem is addressed in [116] (Appendix J) where the standard steady-state Kalman filter is generalized to include a bound on the H_∞ norm of the error signal.

Finally, it is reasonable to expect that in practice both structured and unstructured plant uncertainty will be present. This leads to consideration of the Standard Problem in the presence of parametric uncertainty. Thus it is of interest to design feedback controllers which are guaranteed to satisfy a

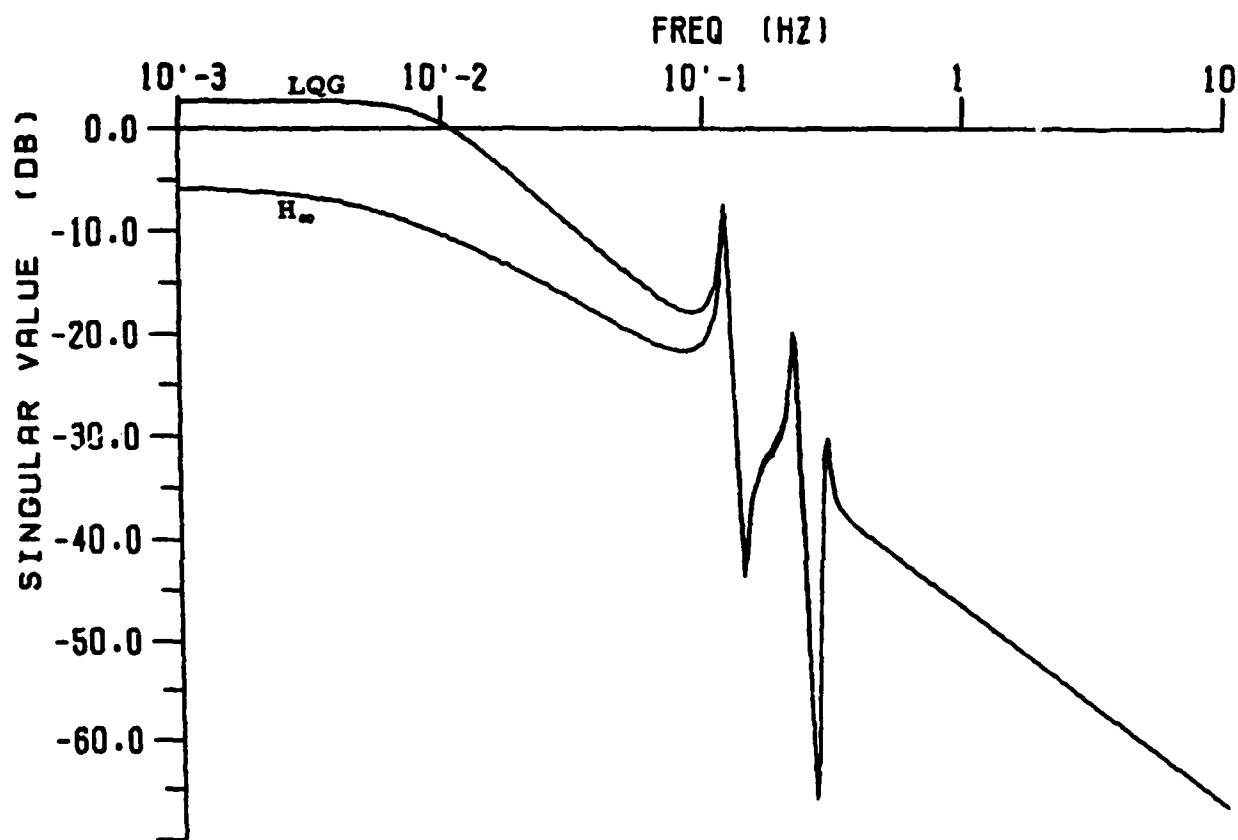


Figure 3-1. The LQG/ H_{∞} Design Equations Yield 9 dB Improvement Over The Corresponding LQG Design for an 8th-Order Nonminimum Phase Plant

specified H_∞ disturbance attenuation constraint over a range of parametric uncertainty. This problem has been addressed in [105] (Appendix J) where the results of [117] on H_∞ design have been merged with these of [94,119] on parametrically robust design. Again the development has been carried out in the context of fixed-order dynamic compensation for maximal design flexibility.

SECTION 4.0
FURTHER EXTENSIONS

4.0 FURTHER EXTENSIONS

4.1 Motivation

The previous sections have addressed two principal problems in control design, namely, fixed-structure design and robustness. Both of these problems concern fundamental issues in the practical implementation of feedback controllers. In this section we extend these results in two directions in order to address larger classes of design problems.

4.2 Tracking

All of the feedback control theory discussed in Sections 2 and 3 addresses the problem of feedback control for regulation in the presence of external disturbances. Many control problems, however, are of a tracking or servomechanism nature. While a limited class of such problems can be recast without loss of generality as regulation problems, many important ones cannot. For example, the standard transformations given in

H. Kwakernaak and R. Sivan, Linear Optimal Control Systems, Wiley, New York, 1972.

B. A. Francis, A Course in H_∞ Control Theory, Springer-Verlag, New York, 1987.

assume that the command signals can be represented as an augmentation of the plant dynamics. There are many important cases, such as the tracking of steps and ramps, which must be represented by uncontrollable, unstable dynamics, where this transformation cannot be applied. Furthermore, such transformations often ignore controller effort. To fill this gap we have undertaken a systematic program for developing a tracking control theory consistent with earlier developments. As a first step we have considered the problem of regulation about a prescribed nonzero set point, which corresponds to the step command tracking problem. Our work in this area was originally motivated by results obtained in

Z. Artstein and A. Leizarowitz, "Tracking Periodic Signals with the Overtaking Criterion," IEEE Trans. Autom. Contr., Vol. AC-30, pp. 1123-1126, 1985.

A. Leizarowitz, "Tracking Nonperiodic Trajectories with the Overtaking Criterion," Appl. Math. Optim., Vol. 14, pp. 155-171, 1986.

A. Leizarowitz, "Infinite Horizon Stochastic Regulation and Tracking With the Overtaking Criterion," Stochastics, 1987.

References [67,103] (Appendix K) present general solutions to the nonzero set point problem for both static and dynamic controllers. The overall controller configurations for these problems are shown in Figures 4-1 and 4-2. Note that these controllers involve two components, namely, a closed-loop feedback component similar to a regulator and an open-loop feedforward component which has no counterpart in the standard theory and which cannot be obtained from standard transformations.

Recent activities have focused on extending the nonzero set point results to broader classes of command and disturbance signals. It turns out that the challenging case (as with steps and ramps) involves signals generated by unstable command or reference dynamics. As a critical first step in addressing this problem we have considered the problem of reduced-order steady-state observer design for unstable plants. These results appear in [125] (Appendix K). This optimal subspace observer problem gives rise to yet another projection which we denote by μ . The most general estimation problem involving all three projections τ , ν , and μ has also been solved and will be reported in [134,139].

4.3 Sampled-Data Control

The discussion in the previous sections has focused on continuous-time systems subject to continuous-time (analog) controllers. In practice, however, controller implementation will almost invariably utilize digital controllers within the context of sampled-data control systems. Rigorous consideration of such systems is critical, particularly for distributed parameter systems which possess modal frequencies beyond the Nyquist rate of any digital controller.

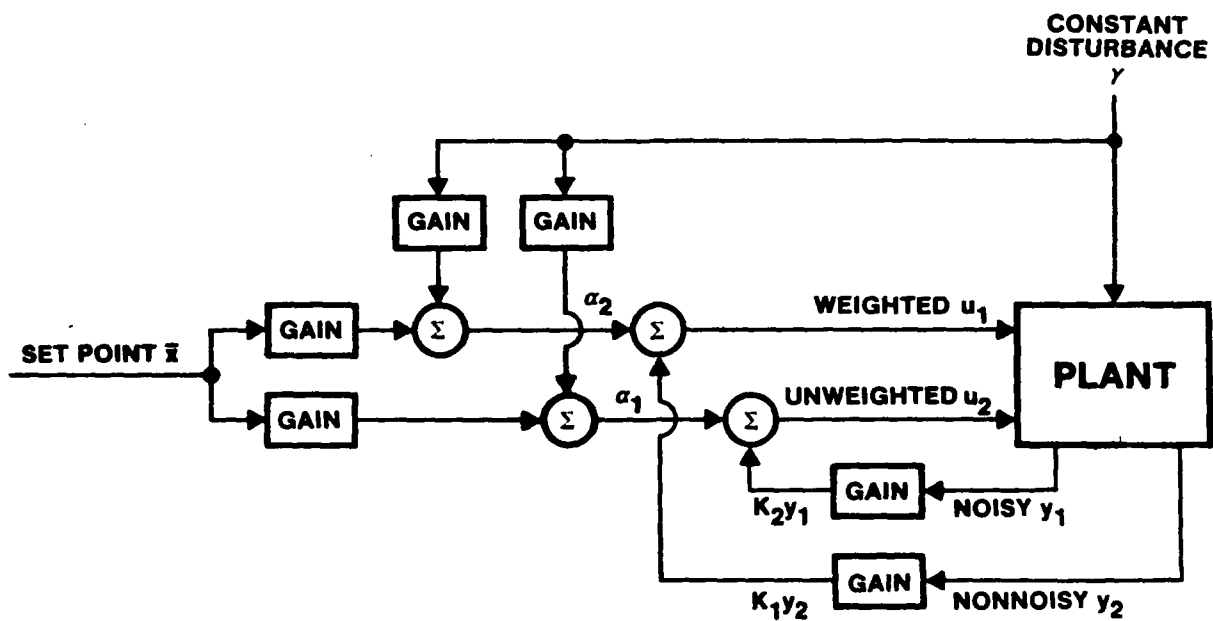
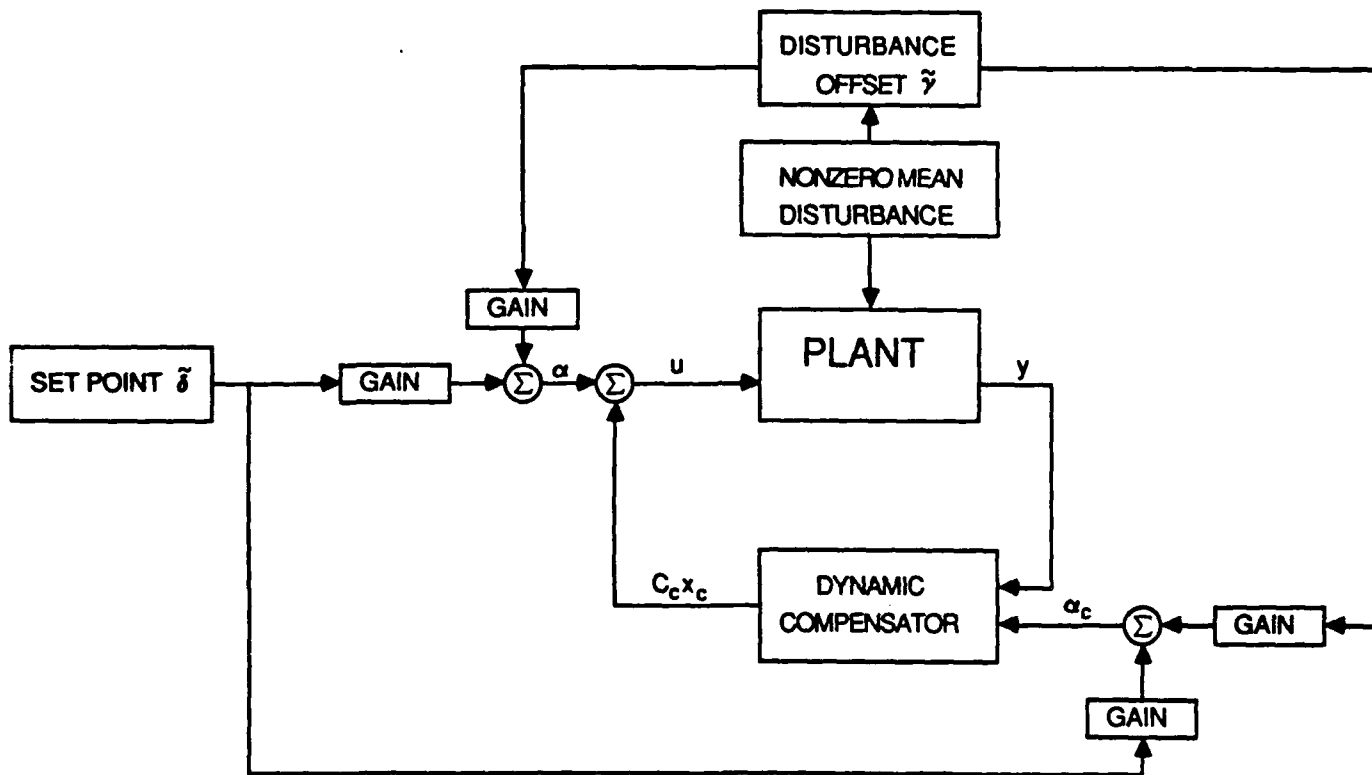


Figure 4-1. The Static Controller for Nonzero Set Point Regulation
Involves Both Feedforward and Feedback Gains



16t

Figure 4-2. The Dynamic Controller for Nonzero Set Point Regulation Involves Both an Open-Loop Gain and a Closed-Loop Dynamic Compensator

Hence, a rigorous theory of sampled-data control design must be developed which accounts precisely for all effects arising from analog-to-digital and digital-to-analog operations.

Optimal projection theory for discrete-time systems was developed in [41] and applied to sampled-data systems in [44] (Appendix L). As a next step it is desirable to obtain robust control results. To this end, the optimal projection equations for reduced-order discrete-time estimation and control in the presence of multiplicative white noise were obtained in [54,69] (Appendix L). After these results were obtained, it became clear that a true sampled-data robustness theory must account for the exponential matrix structure which arises from the sampling process. For example, if $A+\Delta A$ denotes the continuous-time dynamics matrix, where A is the nominal matrix and ΔA denotes uncertainty, then the equivalent discrete-time dynamics matrix is given by $e^{(A+\Delta A)h}$, where h is the sample interval. Because of the exponential function, however, this discrete-time dynamics matrix does not have the additive structure considered in the discrete-time theory in [54,69]. Moreover, a linear approximation for the exponential will not be valid in the presence of system time constants near or above the sample rate.

Although an attempt to bound this discrepancy resulted in new inequalities in [92] and questions of decomposition in [87] (Appendix L), this approach appears inadequate. The crucial clue to the most natural approach was ultimately found in

A. R. Tiedemann and W. L. DeKoning, "The Equivalent Discrete-Time Optimal Control Problem for Continuous-Time Systems with Stochastic Parameters," Int. J. Cont., Vol. 40, pp. 449-466, 1984.

which studied the propagation of multiplicative white noise in the presence of A/D and D/A interfaces. Motivated by these results, we have obtained results which extend the robust performance bounds obtained for continuous-time systems to the sampled-data problem. Specifically, by considering the evolution of the linear parameter uncertainty bound over the sample interval, a robust stability

condition was developed in [128] (Appendix L). This result is unique in that it accounts directly for the exponential structure of the parameter uncertainty.

SECTION 5.0

OPEN PROBLEMS AND FUTURE DIRECTIONS

5.0 OPEN PROBLEMS

5.1 Motivation

The value and importance of the results obtained under this project lie largely in the foundation they provide for future research. The purpose of this section is to collect together various questions and problems as a guide to future endeavors. The order of listing of these questions roughly parallels the order of the previous sections.

5.2 Fixed-Structure Design

Since the fixed-structure design approach involves a nonconvex optimization problem, there arise several questions concerning the structure of the space of solutions.

- Do there exist verifiable a priori conditions which guarantee stabilizability of a given linear time-invariant plant by fixed-order dynamic compensation? As in the full-order case, one would expect such conditions to play a fundamental role in determining the existence of solutions to the design equations. Conversely, when the plant is known to be stabilizable by a controller of order n_c , does the underlying optimization problem always possess a solution? Will the design equations always yield at least one such stabilizing controller? How is the ability to find stabilizing controllers affected by the choice of weightings and noise intensities?
- Is it possible to design all subcontrollers of a decentralized dynamic compensator simultaneously without performing sequential iterations? If a sequential algorithm is used, then under what conditions is the algorithm guaranteed to find the global minimum?
- How can the fixed-structure approach be extended to address the simultaneous stabilization problem, i.e., the problem of finding a single controller which stabilizes several different plants simultaneously?
- The L_2 model reduction theory of [32] (Appendix B) can readily be extended to the problem of characterizing optimal finite-dimensional models for infinite-dimensional systems using the method of [37] (Appendix C). Can such finite-

dimensional models serve as useful lumped approximations to distributed parameter systems? Can the L_2/H_∞ model reduction theory of [114] (Appendix J) be used similarly?

- How can the fixed-structure approach be used to design controllers with additional constraints on their internal structure, such as prespecified pole locations? This question is the basis for ongoing work in [131].

5.3 Robust Analysis and Design

There exist a variety of open questions concerning the conservatism and effectiveness of the parametric robustness bounds and the H_∞ design equations.

- For which class of parameter uncertainty structures are the quadratic Lyapunov bounds nonconservative? How can the robustified design equations be used iteratively to reduce design conservatism?
- The multiplicative noise model was shown in [77] (Appendix F) to guarantee deterministic robustness. However, this result involved a uniform right shift rather than the variable left shift arising from the Stratonovich interpretation of the multiplicative noise. Can it be shown rigorously that the Stratonovich model yields robust controllers? Furthermore, can the relationship between Stratonovich design and positive real controllers for modal systems be made precise?
- The basis for the H_∞ design results obtained in [117] (Appendix J) is the quadratic bound developed for parametrically robust control in [94] (Appendix I). This raises the following question: Does there exist an alternative interpretation of the linear bound which can be used to guarantee disturbance attenuation for some specified class of disturbances?
- The H_∞ control design results are virtually identical to the optimality conditions for the problem of minimizing an exponential-of-quadratic cost criterion as considered in

P. R. Kumar and J. H. van Schuppen, "On the Optimal Control of Stochastic Systems With an Exponential-of-Integral Performance Index," J. Math. Anal. Appl., Vol. 80, pp. 312-332, 1981.

P. Whittle, "Risk-Sensitive Linear/Quadratic/Gaussian Control," Adv. Appl. Prob., Vol. 13, pp. 764-777, 1981.

A. Bensoussan and J. H. van Schuppen, "Optimal Control of Partially Observable Stochastic Systems With an Exponential-of-Integral Performance Index," SIAM J. Contr. Optim., Vol. 23, pp. 599-613, 1985.

P. Whittle and J. Kuhn, "A Hamiltonian Formulation of Risk-Sensitive Linear/Quadratic/Gaussian Control," Int. J. Contr. Vol. 43, pp. 1-12, 1986.

Is it possible to directly extend these results using the fixed structure approach? Also, can the fixed-structure approach be used to extend the Maximum Entropy theory of

D. Mustafa and K. Glover, "Controllers Which Satisfy a Closed-Loop H_∞ Norm Bound and Maximize an Entropy Integral," Proc. IEEE Conf. Dec. Contr., Austin, TX, December 1988.

- The L_2/H_∞ model reduction theory given in [114] (Appendix J) minimizes an L_2 criterion subject to a constraint on the H_∞ distance between the full- and reduced-order models. Can the L_2 criterion be neutralized so as to obtain a "pure" H_∞ result as is done in [117] (Appendix J) for full-order control design? Can the resulting H_∞ solution be shown to actually characterize the H_∞ optimal reduced-order model by taking the H_∞ constraint to be sufficiently small? Similar questions apply to fixed-order control design. For example, does there exist a "pure" H_∞ reduced-order control design theory? Can these results be shown to be necessary as well as sufficient?
- What is the generalization of the H_∞ control and estimation results to the singular problem? To the cross-weighting problem?
- Is it possible to extend the L_2 and L_2/H_∞ model reduction results to allow the reduced-order model to be nonstrictly proper?

5.4 Tracking and Sampled-Data Control

With regard to tracking and sampled-data theory a number of problems remain to be explored.

- Is it possible to develop a methodology for designing tracking controllers which applies to a broad range of signal models? For example, the command signal may be known exactly in advance (such as a specified square wave) while, at the other extreme, it may only be known to be an element of a large class of signals. For example, step commands are known to be steps but their exact level is not known until they actually occur during operation. Other command signals may only be known to be outputs of systems driven by random noise. A classification scheme based upon the degree and type of priori knowledge of the command signal should lead to a hierarchy of control designs ranging from poorly known to well-known command signals. In addition, it is important to distinguish between a priori command signal knowledge available during the design phase and command signal knowledge available during operation. The differences between these cases can be used to account for differing assumptions appearing in the literature. Relevant references include

B. D. O. Anderson and J. B. Moore, Linear Optimal Control, Prentice-Hall, Englewood Cliffs, NJ, 1970.

C. D. Johnson, "Accommodation of External Disturbances in Linear Regulator and Servomechanism Problems," IEEE Trans. Autom. Contr., Vol. AC-16, pp. 635-644, 1971.

E. J. Davison and A. Goldenberg, "Robust Control of a General Servomechanism Problem: The Servo Compensator," Automatica, Vol. 11, pp. 461-471, 1975.

E. J. Davison, "The Robust Decentralized Control of a General Servomechanism Problem," IEEE Trans. Autom. Contr., Vol. AC-21, pp. 14-24, 1976.

E. J. Davison, "The Robust Control of a Servomechanism Problem for Linear Time-Invariant Multivariable Systems," IEEE Trans. Autom. Contr., Vol. AC-21, pp. 25-34, 1976.

E. J. Davison, "Multivariable Tuning Regulators: The Feedforward and Robust Control of a General Servomechanism Problem," IEEE Trans. Autom. Contr., Vol. AC-21, pp. 35-47, 1976.

C. A. Desoer and Y. T. Wang, "Linear Time-Invariant Robust Servomechanism Problem: A Self-Contained Exposition," in Control and Dynamic Systems, Vol. 16, C. T. Leondes, Ed., pp. 81-129, Academic Press, New York, 1980.

E. J. Davison and I. J. Ferguson, "The Design of Controllers for the Multivariable Robust Servomechanism Problem Using Parameter Optimization Methods," IEEE Trans. Autom. Contr., Vol. AC-26, pp. 93-110, 1981.

J. D. Turner, H. M. Chun and J.-N. Juang, "Closed-Form Solutions for a Class of Optimal Quadratic Tracking Problems," J. Optim. Thy. Appl., Vol. 47, pp. 465-481, 1985.

J.-N. Juang, J. D. Turner and H. M. Chun, "Closed-Form Solutions for Feedback Control with Terminal Constraints," AIAA J. Guid. Contr. Dyn., Vol. 8, pp. 39-43, 1985.

W. E. Schmitendorf and B. R. Barmish, "Robust Asymptotic Tracking for Linear Systems with Unknown Parameters," Automatica, Vol. 22, pp. 355-360, 1986.

- How can the new subspace projection μ , which arises in the observer design problem in [125] (Appendix K), be used to design servocompensators? That is, can μ be used to design controllers which track the output of an unstable command model?
- Is it possible to develop a theory of robust sampled-data controller synthesis which accounts directly for the exponential structure of the equivalent discrete-time model? The results of [128] (Appendix L) provide a starting point in this regard.
- What is the form of the equations for the H_∞ -constrained discrete-time control-design problem?

SECTION 6.0
COMPREHENSIVE REFERENCE LIST

1. D. C. Hyland, "The Modal Coordinate/Radiative Transfer Formulation of Structural Dynamics--Implications for Vibration Suppression in Large Space Platforms," MIT Lincoln Laboratory, TR-27, 14 March 1979.
2. D. C. Hyland, "Optimal Regulation of Structural Systems With Uncertain Parameters," MIT Lincoln Laboratory, TR-551, 2 February 1981, DDC# ADA-099111/7.
3. D. C. Hyland, "Active Control of Large Flexible Spacecraft: A New Design Approach Based on Minimum Information Modelling of Parameter Uncertainties," Proc. Third VPI&SU/AIAA Symposium, pp. 631-646, Blacksburg, VA, June 1981.
4. D. C. Hyland, "Optimal Regulator Design Using Minimum Information Modelling of Parameter Uncertainties: Ramifications of the New Design Approach," Proc. Third VPI&SU/AIAA Symposium, pp. 701-716, Blacksburg, VA, June 1981.
5. D. C. Hyland and A. N. Madiwale, "Minimum Information Approach to Regulator Design: Numerical Methods and Illustrative Results," Proc. Third VPI&SU/AIAA Symposium, pp. 101-118, Blacksburg, VA, June 1981.
6. D. C. Hyland and A. N. Madiwale, "A Stochastic Design Approach for Full-Order Compensation of Structural Systems with Uncertain Parameters," Proc. AIAA Guid. Contr. Conf., pp. 324-332, Albuquerque, NM, August 1981.
7. D. C. Hyland, "Optimality Conditions for Fixed-Order Dynamic Compensation of Flexible Spacecraft with Uncertain Parameters," AIAA 20th Aerospace Sciences Meeting, paper 82-0312, Orlando, FL, January 1982.
8. D. C. Hyland, "Structural Modeling and Control Design Under Incomplete Parameter Information: The Maximum Entropy Approach," AFOSR/NASA Workshop on Modeling, Analysis and Optimization Issues for Large Space Structures, Williamsburg, VA, May 1982.
9. D. C. Hyland, "Minimum Information Stochastic Modelling of Linear Systems with a Class of Parameter Uncertainties," Proc. Amer. Contr. Conf., pp. 620-627, Arlington, VA, June 1982.
10. D. C. Hyland, "Maximum Entropy Stochastic Approach to Control Design for Uncertain Structural Systems," Proc. Amer. Contr. Conf., pp. 680-688, Arlington, VA, June 1982.
11. D. C. Hyland, "Minimum Information Modeling of Structural Systems with Uncertain Parameters," Proceedings of the Workshop on Applications of Distributed System Theory to the Control of Large Space Structures, G. Rodriguez, ed., pp. 71-88, JPL, Pasadena, CA, July 1982.
12. D. C. Hyland and A. N. Madiwale, "Fixed-Order Dynamic Compensation Through Optimal Projection," Proceedings of the Workshop on Applications of Distributed System Theory to the Control of Large Space Structures, G. Rodriguez, ed., pp. 409-427, JPL, Pasadena, CA, July 1982.

13. D. C. Hyland, "Mean-Square Optimal Fixed-Order Compensation--Beyond Spillover Suppression," paper 1403, AIAA Astrodynamics Conference, San Diego, CA, August 1982.
14. D. C. Hyland, "Robust Spacecraft Control Design in the Presence of Sensor/Actuator Placement Errors," AIAA Astrodynamics Conference, San Diego, CA, August 1982.
15. D. C. Hyland, "The Optimal Projection Approach to Fixed-Order Compensation: Numerical Methods and Illustrative Results," AIAA 21st Aerospace Sciences Meeting, paper 83-0303, Reno, NV, January 1983.
16. D. C. Hyland, "Mean-Square Optimal, Full-Order Compensation of Structural Systems with Uncertain Parameters," MIT Lincoln Laboratory, TR-626, 1 June 1983.
17. D. S. Bernstein and D. C. Hyland, "Explicit Optimality Conditions for Finite-Dimensional Fixed-Order Dynamic Compensation of Infinite-Dimensional Systems," presented at SIAM Fall Meeting, Norfolk, VA, November 1983.
18. D. C. Hyland and D. S. Bernstein, "Explicit Optimality Conditions for Fixed-Order Dynamic Compensation," Proc. 22nd IEEE Conf. Dec. Contr., pp. 161-165, San Antonio, TX, December 1983.
19. F. M. Ham, J. W. Shipley and D. C. Hyland, "Design of a Large Space Structure Vibration Control Experiment," Proc. 2nd Int. Modal Anal. Conf., pp. 550-558, Orlando, FL, February 1984.
20. D. C. Hyland, "Comparison of Various Controller-Reduction Methods: Suboptimal Versus Optimal Projection," Proc. AIAA Dynamics Specialists Conf., pp. 381-389, Palm Springs, CA, May 1984.
21. D. S. Bernstein and D. C. Hyland, "The Optimal Projection Equations for Fixed-Order Dynamic Compensation of Distributed Parameter Systems," Proc. AIAA Dynamics Specialists Conf., pp. 396-400, Palm Springs, CA, May 1984.
22. F. M. Ham and D. C. Hyland, "Vibration Control Experiment Design for the 15-M Hoop/Column Antenna," Proceedings of the Workshop on the Identification and Control of Flexible Space Structures, pp. 229-252, San Diego, CA, June 1984.
23. D. S. Bernstein and D. C. Hyland, "Numerical Solution of the Optimal Model Reduction Equations," Proc. AIAA Guid. Contr. Conf., pp. 560-562, Seattle, WA, August 1984.
24. D. C. Hyland and D. S. Bernstein, "The Optimal Projection Equations for Fixed-Order Dynamic Compensation," IEEE Trans. Autom. Contr., Vol. AC-29, pp. 1034-1037, 1984.

25. D. C. Hyland, "Application of the Maximum Entropy/Optimal Projection Control Design Approach for Large Space Structures," Proc. Large Space Antenna Systems Technology Conference, pp. 617-654, NASA Langley, December 1984.
26. D. C. Hyland and D. S. Bernstein, "The Optimal Projection Approach to Model Reduction and the Relationship Between the Methods of Wilson and Moore," Proc. 23rd IEEE Conf. Dec. Contr., pp. 120-126, Las Vegas, NV, December 1984.
27. D. S. Bernstein and D. C. Hyland, "The Optimal Projection Approach to Designing Optimal Finite-Dimensional Controllers for Distributed Parameter Systems," Proc. 23rd IEEE Conf. Dec. Contr., pp. 556-560, Las Vegas, NV, December 1984.
28. L. D. Davis, D. C. Hyland and D. S. Bernstein, "Application of the Maximum Entropy Design Approach to the Spacecraft Control Laboratory Experiment (SCOLE)," Final Report, NASA Langley, January 1985.
29. D. S. Bernstein and D. C. Hyland, "The Optimal Projection Equations for Reduced-Order State Estimation," IEEE Trans. Autom. Contr., Vol. AC-30, pp. 583-585, 1985.
30. D. S. Bernstein and D. C. Hyland, "The Optimal Projection Equations for Reduced-Order State Estimation," Proc. Amer. Contr. Conf., pp. 164-167, Boston, MA, June 1985.
31. D. S. Bernstein and D. C. Hyland, "Optimal Projection/Maximum Entropy Stochastic Modelling and Reduced-Order Design Synthesis," Proc. IFAC Workshop on Model Error Concepts and Compensation, Boston, MA, June 1985, R. E. Skelton and D. H. Owens, Eds., pp. 47-54, Pergamon Press, Oxford, 1986.
32. D. C. Hyland and D. S. Bernstein, "The Optimal Projection Equations for Model Reduction and the Relationships Among the Methods of Wilson, Skelton and Moore," IEEE Trans. Autom. Contr., Vol. AC-30, pp. 1201-1211, 1985.
33. D. S. Bernstein, "The Optimal Projection Equations for Fixed-Structure Decentralized Dynamic Compensation," Proc. 24th IEEE Conf. Dec. Contr., pp. 104-107, Fort Lauderdale, FL, December 1985.
34. D. S. Bernstein, L. D. Davis, S. W. Greeley and D. C. Hyland, "The Optimal Projection Equations for Reduced-Order, Discrete-Time Modelling, Estimation and Control," Proc. 24th IEEE Conf. Dec. Contr., pp. 573-578, Fort Lauderdale, FL, December 1985.
35. D. S. Bernstein and D. C. Hyland, "The Optimal Projection/Maximum Entropy Approach to Designing Low-Order, Robust Controllers for Flexible Structures," Proc. 24th IEEE Conf. Dec. Contr., pp. 745-752, Fort Lauderdale, FL, December 1985.

36. D. S. Bernstein, L. D. Davis, S. W. Greeley and D. C. Hyland, "Numerical Solution of the Optimal Projection/Maximum Entropy Design Equations for Low-Order, Robust Controller Design," Proc. 24th IEEE Conf. Dec. Contr., pp. 1795-1798, Fort Lauderdale, FL, December 1985.
37. D. S. Bernstein and D. C. Hyland, "The Optimal Projection Equations for Finite-Dimensional Fixed-Order Dynamic Compensation of Infinite-Dimensional Systems," SIAM J. Contr. Optim., Vol. 24, pp. 122-151, 1986.
38. J. W. Shipley and D. C. Hyland, "The Mast Flight System Dynamic Characteristics and Actuator/Sensor Selection and Location," Proc. 9th Annual AAS Guid. Contr. Conf., Keystone, CO, February 1986, R. D. Culp and J. C. Durrett, Eds., pp. 31-49, American Astronautical Society, San Diego, CA, 1986.
39. D. S. Bernstein and S. W. Greeley, "Robust Controller Synthesis Using the Maximum Entropy Design Equations," IEEE Trans. Autom. Contr., Vol. AC-31, pp. 362-364, 1986.
40. D. C. Hyland, D. S. Bernstein, L. D. Davis, S. W. Greeley and S. Richter, "MEOP: Maximum Entropy/Optimal Projection Stochastic Modelling and Reduced-Order Design Synthesis," Final Report, Air Force Office of Scientific Research, Bolling AFB, Washington, DC, April 1986.
41. D. S. Bernstein, L. D. Davis and D. C. Hyland, "The Optimal Projection Equations for Reduced-Order, Discrete-Time Modelling, Estimation and Control," AIAA J. Guid. Contr. Dyn., Vol. 9, pp. 288-293, 1986.
42. D. S. Bernstein, L. D. Davis and S. W. Greeley, "The Optimal Projection Equations for Fixed-Order, Sampled-Data Dynamic Compensation with Computation Delay," Proc. Amer. Contr. Conf., pp. 1590-1597, Seattle, WA, June 1986.
43. D. S. Bernstein and S. W. Greeley, "Robust Output-Feedback Stabilization: Deterministic and Stochastic Perspectives," Proc. Amer. Contr. Conf., pp. 1818-1826, Seattle, WA, June 1986.
44. D. S. Bernstein, L. D. Davis and S. W. Greeley, "The Optimal Projection Equations for Fixed-Order, Sampled-Data Dynamic Compensation with Computation Delay," IEEE Trans. Autom. Contr., Vol. AC-31, pp. 859-862, 1986.
45. D. S. Bernstein, "Optimal Projection/Guaranteed Cost Control Design Synthesis: Robust Performance via Fixed-Order Dynamic Compensation," presented at SIAM Conference on Linear Algebra in Signals, Systems and Control, Boston, MA, August 1986.
46. D. C. Hyland, "Control Design Under Stratonovich Models: Robust Stability Guarantees via Lyapunov Matrix Functions," presented at SIAM Conference on Linear Algebra in Signals, Systems and Control, Boston, MA, August 1986.

47. D. S. Bernstein, "OFUS: Optimal Projection for Uncertain Systems," Annual Report, Air Force Office of Scientific Research, Bolling AFB, Washington, DC, October 1986.
48. B. J. Boan and D. C. Hyland, "The Role of Metal Matrix Composites for Vibration Suppression in Large Space Structures," Proc. MMC Spacecraft Survivability Tech. Conf., MMCIAC Kaman Tempo Publ., Stanford Research Institute, Palo Alto, CA, October 1986.
49. L. D. Davis, T. Otten, F. M. Ham and D. C. Hyland, "Mast Flight System Dynamic Performance," presented at 1st NASA/DOD CSI Technology Conference, Norfolk, VA, November 1986.
50. D. C. Hyland, "An Experimental Testbed for Validation of Control Methodologies in Large Space Optical Structures," in Structural Mechanics of Optical Systems II, pp. 146-155, A. E. Hatheway, ed., Proceedings of SPIE, Vol. 748, Optoelectronics and Laser Applications Conference, Los Angeles, CA, January 1987.
51. J. W. Shipley, L. D. Davis, W. T. Burton and F. M. Ham, "Development of the Mast Flight System Linear DC Motor Inertial Actuator," Proc. 10th Annual AAS Guid. Contr. Conf., Keystone, CO, February 1987, R. D. Culp and T. J. Kelley, Eds., pp. 237-255, American Astronautical Society, San Diego, CA, 1987.
52. W. M. Haddad, Robust Optimal Projection Control-System Synthesis, Ph.D. Dissertation, Department of Mechanical Engineering, Florida Institute of Technology, Melbourne, FL, March 1987.
53. A. W. Daubendiek and R. G. Brown, "A Robust Kalman Filter Design," Report ISU-ERI-Ames-87226, College of Engineering, Iowa State University, Ames, IA, March 1987.
54. W. M. Haddad and D. S. Bernstein, "The Optimal Projection Equations for Discrete-Time Reduced-Order State Estimation for Linear Systems with Multiplicative White Noise," Sys. Contr. Lett., Vol. 8, pp. 381-388, 1987.
55. D. C. Hyland and D. S. Bernstein, "MEOP Control Design Synthesis: Optimal Quantification of the Major Design Tradeoffs," in Structural Dynamics and Control Interaction of Flexible Structures, Part 2, pp. 1033-1070, NASA Conf. Publ. 2467, 1987.
56. F. M. Ham and S. W. Greeley, "Active Damping Control Design for the Mast Flight System," Proc. Amer. Contr. Conf., pp. 355-367, Minneapolis, MN, June 1987.
57. D. C. Hyland and D. S. Bernstein, "Uncertainty Characterization Schemes: Relationships to Robustness Analysis and Design," presented at Amer. Contr. Conf., Minneapolis, MN, June 1987.
58. W. M. Haddad and D. S. Bernstein, "The Optimal Projection Equations for Reduced-Order State Estimation: The Singular Measurement Noise Case," Proc. Amer. Contr. Conf., pp. 779-785, Minneapolis, MN, June 1987.

59. D. C. Hyland and D. S. Bernstein, "The Majorant Lyapunov Equation: A Nonnegative Matrix Equation for Robust Stability and Performance of Large Scale Systems," Proc. Amer. Contr. Conf., pp. 910-917, Minneapolis, MN, June 1987.
60. D. S. Bernstein, "Sequential Design of Decentralized Dynamic Compensators Using the Optimal Projection Equations: An Illustrative Example Involving Interconnected Flexible Beams," Proc. Amer. Contr. Conf., pp. 986-989, Minneapolis, MN, June 1987.
61. L. D. Davis, "Issues in Sampled-Data Control: Discretization Cost and Aliasing," presented at Amer. Contr. Conf., Minneapolis, MN, June 1987.
62. D. C. Hyland, "Majorant Bounds, Stratonovich Models, and Statistical Energy Analysis," presented at Amer. Contr. Conf., Minneapolis, MN, June 1987.
63. S. Richter, "A Homotopy Algorithm for Solving the Optimal Projection Equations for Fixed-Order Dynamic Compensation: Existence, Convergence and Global Optimality," Proc. Amer. Contr. Conf., pp. 1527-1531, Minneapolis, MN, June 1987.
64. D. S. Bernstein, "The Optimal Projection Equations For Nonstrictly Proper Fixed-Order Dynamic Compensation," Proc. Amer. Contr. Conf., pp. 1991-1996, Minneapolis, MN, June 1987.
65. D. S. Bernstein and W. M. Haddad, "Optimal Output Feedback for Nonzero Set Point Regulation," Proc. Amer. Contr. Conf., pp. 1997-2003, Minneapolis, MN, June 1987.
66. D. S. Bernstein and S. Richter, "A Homotopy Algorithm for Solving the Optimal Projection Equations for Fixed-Order Dynamic Compensation," presented at Int. Symp. Math. Thy. Net. Sys., Phoenix, AZ, June 1987.
67. D. S. Bernstein and W. M. Haddad, "Optimal Output Feedback for Nonzero Set Point Regulation," IEEE Trans. Autom. Contr., Vol. AC-32, pp. 641-645, 1987.
68. S. Richter, "Reduced-Order Control Design via the Optimal Projection Approach: A Homotopy Algorithm for Global Optimality," Proc. Sixth VPI&SU Symp. Dyn. Contr. Large Str., L. Meirovitch, Ed., pp. 17-30, Blacksburg, VA, June 1987.
69. D. S. Bernstein and W. M. Haddad, "Optimal Projection Equations for Discrete-Time Fixed-Order Dynamic Compensation of Linear Systems with Multiplicative White Noise," Int. J. Contr., Vol. 46, pp. 65-73, 1987.
70. F. M. Ham, B. L. Henniges and S. W. Greeley, "Active Damping Control Design for the COFS Mast Flight System," Proc. AIAA Conf. Guid. Nav. Contr., pp. 354-360, Monterey, CA, August 1987.
71. S. W. Greeley and D. C. Hyland, "Reduced-Order Compensation: LQG Reduction Versus Optimal Projection," Proc. AIAA Conf. Guid. Nav. Contr., pp. 605-616, Monterey, CA, August 1987.

72. D. S. Bernstein, "Decentralized Control of Large Space Structures via Fixed-Order Dynamic Compensation," presented at IEEE CSS Workshop on Current Issues in Decentralized and Distributed Control, Columbus, OH, September 1987.
73. D. S. Bernstein, "OFUS: Optimal Projection for Uncertain Systems," Annual Report, Air Force Office of Scientific Research, Bolling AFB, Washington, DC, October 1987.
74. R. C. Talcott, J. W. Shipley, T. Kimball and S. W. Greeley, "Mast Flight System Engineering Development and System Integration," presented at 2nd NASA/DOD CSI Technology Conference, Colorado Springs, CO, November 1987.
75. D. C. Hyland and D. S. Bernstein, "The Majorant Lyapunov Equation: A Nonnegative Matrix Equation for Guaranteed Robust Stability and Performance of Large Scale Systems," IEEE Trans. Autom. Contr., Vol. AC-32, pp. 1005-1013, 1987.
76. D. S. Bernstein, "Sequential Design of Decentralized Dynamic Compensators Using the Optimal Projection Equations," Int. J. Contr., Vol. 46, pp. 1569-1577, 1987.
77. D. S. Bernstein, "Robust Static and Dynamic Output-Feedback Stabilization: Deterministic and Stochastic Perspectives," IEEE Trans. Autom. Contr., Vol. AC-32, pp. 1076-1084, 1987.
78. W. M. Haddad and D. S. Bernstein, "The Optimal Projection Equations for Reduced-Order State Estimation: The Singular Measurement Noise Case," IEEE Trans. Autom. Contr., Vol. AC-32, pp. 1135-1139, 1987.
79. D. S. Bernstein, "The Optimal Projection Equations for Static and Dynamic Output Feedback: The Singular Case," IEEE Trans. Autom. Contr., Vol. AC-32, pp. 1139-1143, 1987.
80. W. M. Haddad and D. S. Bernstein, "The Unified Optimal Projection Equations for Simultaneous Reduced-Order, Robust Modeling, Estimation and Control," Proc. IEEE Conf. Dec. Contr., pp. 449-454, Los Angeles, CA, December 1987.
81. S. W. Greeley and D. C. Hyland, "Reduced-Order Compensation: IQG Reduction Versus Optimal Projection Using a Homotopic Continuation Method," Proc. IEEE Conf. Dec. Contr., pp. 742-747, Los Angeles, CA, December 1987.
82. D. S. Bernstein and W. M. Haddad, "The Optimal Projection Equations with Petersen-Hollot Bounds: Robust Controller Synthesis with Guaranteed Structured Stability Radius," Proc. IEEE Conf. Dec. Contr., pp. 1308-1318, Los Angeles, CA, December 1987.
83. W. M. Haddad and D. S. Bernstein, "Robust, Reduced-Order, Nonstrictly Proper State Estimation via the Optimal Projection Equations with Petersen-Hollot Bounds," Sys. Contr. Lett., Vol. 9, pp. 423-431, 1987.

84. D. C. Hyland, "Experimental Investigations in Active Vibration Control for Application to Large Space Systems," in Space Structures, Power, and Power Conditioning, R. F. Askew, Ed., Proc. SPIE, Vol. 871, pp. 242-253, 1988.
85. D. C. Hyland and D. J. Phillips, "Development of the Linear Precision Actuator," 11th Annual AAS Guid. Contr. Conf., Keystone, CO, January 1988.
86. W. M. Haddad and D. S. Bernstein, "Optimal Output Feedback for Nonzero Set Point Regulation: The Discrete-Time Case," Int. J. Contr., Vol. 47, pp. 529-536, 1988.
87. D. S. Bernstein, "Commuting Matrix Exponentials," Problem 88-1, SIAM Review, Vol. 30, p. 123, 1988.
88. D. S. Bernstein and D. C. Hyland, "Optimal Projection for Uncertain Systems (OPUS): A Unified Theory of Reduced-Order, Robust Control Design," in Large Space Structures: Dynamics and Control, S. N. Atluri and A. K. Amos, Eds., pp. 263-302, Springer-Verlag, New York, 1988.
89. W. M. Haddad and D. S. Bernstein, "The Unified Optimal Projection Equations for Simultaneous Reduced-Order, Robust Modeling, Estimation and Control," Int. J. Contr., Vol. 47, pp. 1117-1132, 1988.
90. D. C. Hyland, "Homotopic Continuation Methods for the Design of Optimal Fixed-Form Dynamic Controllers," in Computational Mechanics '85, Proc. Int. Conf. Comp. Eng. Sci., S. N. Atluri and G. Yagawa, Eds., Vol. 2, pp. 44.i.1-4, 1988.
91. D. C. Hyland and J. W. Shipley, "A Unified Process for System Identification Based on Performance Assessment," in Model Determination for Large Space Systems Workshop, pp. 570-595, Pasadena, CA, March 1988.
92. D. S. Bernstein, "Inequalities for the Trace of Matrix Exponentials," SIAM J. Matrix Anal. Appl., Vol. 9, pp. 156-158, 1988.
93. D. C. Hyland, D. S. Bernstein, and E. G. Collins, Jr., "Maximum Entropy/Optimal Projection Design Synthesis for Decentralized Control of Large Space Structures," Final Report, Air Force Office of Scientific Research, Bolling AFB, Washington, DC, May 1988.
94. D. S. Bernstein and W. M. Haddad, "The Optimal Projection Equations with Petersen-Hollot Bounds: Robust Stability and Performance via Fixed-Order Dynamic Compensation for Systems with Structured Real-Valued Parameter Uncertainty," IEEE Trans. Autom. Contr., Vol. 33, pp. 578-582, 1988.
95. W. M. Haddad and D. S. Bernstein, "Robust, Reduced-Order, Nonstrictly Proper State Estimation via the Optimal Projection Equations with Guaranteed Cost Bounds," IEEE Trans. Autom. Contr., Vol. 33, pp. 591-595, 1988.

96. D. S. Bernstein and W. M. Haddad, "LQG Control with an H_∞ Performance Bound: A Riccati Equation Approach," Proc. Amer. Contr. Conf., pp. 796-802, Atlanta, GA, June 1988.
97. S. W. Greeley, D. J. Phillips and D. C. Hyland, "Experimental Demonstration of Maximum Entropy, Optimal Projection Design Theory for Active Vibration Control," Proc. Amer. Contr. Conf., pp. 1462-1467, Atlanta, GA, June 1988.
98. D. C. Hyland and E. G. Collins, Jr., "A Robust Control Experiment Using an Optical Structure Prototype," Proc. Amer. Contr. Conf., pp. 2046-2049, Atlanta, GA, June 1988.
99. D. S. Bernstein and W. M. Haddad, "Robust Stability and Performance for Fixed-Order Dynamic Compensation via the Optimal Projection Equations with Guaranteed Cost Bounds," Proc. Amer. Contr. Conf., pp. 2471-2476, Atlanta, GA, June 1988.
100. S. Richter, "Recent Advances in the Application of Homotopic Continuation Methods to Control Problems," presented at Amer. Contr. Conf., Atlanta, GA, June 1988.
101. W. M. Haddad and D. S. Bernstein, "Robust, Reduced-Order Modeling via the Optimal Projection Equations with Petersen-Hollot Bounds," IEEE Trans. Autom. Contr., Vol. 33, pp. 692-695, 1988.
102. S. W. Greeley and D. C. Hyland, "Reduced-Order Compensation: Linear-Quadratic Reduction Versus Optimal Projection," AIAA J. Guid. Contr. Dyn., Vol. 11, pp. 328-335, 1988.
103. W. M. Haddad and D. S. Bernstein, "Optimal Nonzero Set Point Regulation via Fixed-Order Dynamic Compensation," IEEE Trans. Autom. Contr., Vol. 33, pp. 848-852, 1988.
104. D. S. Bernstein and D. C. Hyland, "Optimal Projection Equations for Reduced-Order Modelling, Estimation and Control of Linear Systems with Multiplicative White Noise," J. Optim. Thy. Appl., Vol. 58, pp. 387-409, 1988.
105. A. N. Madiwale, W. M. Haddad, and D. S. Bernstein, "Robust H_∞ Controllers for Systems with Parameter Uncertainty," Proc. IEEE Conf. Dec. Contr., Austin, TX, December 1988.
106. D. S. Bernstein and W. M. Haddad, "Robust Decentralized Output Feedback: The Static Controller Case," Proc. IEEE Conf. Dec. Contr., Austin, TX, December 1988.
107. D. S. Bernstein and I. G. Rosen, "An Approximation Technique for Computing Optimal Fixed-Order Controllers for Infinite-Dimensional Systems," Proc. IEEE Conf. Dec. Contr., Austin, TX, December 1988.
108. D. C. Hyland and E. G. Collins, Jr., "An M-Matrix and Majorant Approach to Robust Stability and Performance Analysis for Systems with Structured Uncertainty," Proc. IEEE Conf. Dec. Contr., Austin, TX, December 1988.

109. D. S. Bernstein and W. M. Haddad, "Robust Stability and Performance Analysis for State Space Systems via Quadratic Lyapunov Bounds," Proc. IEEE Conf. Dec. Contr., Austin, TX, December 1988.
110. E. G. Collins, Jr., and D. C. Hyland, "Improved Robust Performance Bounds in Covariance Majorant Analysis," Proc. IEEE Conf. Dec. Contr., Austin, TX, December 1988.
111. D. S. Bernstein and W. M. Haddad, "Optimal Reduced-Order State Estimation for Unstable Plants," Proc. IEEE Conf. Dec. Contr., Austin, TX, December 1988.

To Appear

- 112. D. C. Hyland and E. G. Collins, Jr., "Block Kronecker Products and Block-Norm Matrices in Large Scale Systems Analysis," SIAM J. Matrix Anal. Appl., to appear.
- 113. D. S. Bernstein, "Robust Stability and Performance via Fixed-Order Dynamic Compensation," SIAM J. Contr. Optim., to appear.
- 114. W. M. Haddad and D. S. Bernstein, "Combined L_2/H_∞ Model Reduction," Int. J. Contr., to appear.
- 115. D. S. Bernstein and W. M. Haddad, "Robust Stability and Performance Analysis for Linear Dynamic Systems," IEEE Trans. Autom. Contr., Vol. 34, 1989, to appear.
- 116. D. S. Bernstein and W. M. Haddad, "Steady-State Kalman Filtering with an H_∞ Error Bound," Sys. Contr. Lett., Vol. 12, 1989, to appear.
- 117. D. S. Bernstein and W. M. Haddad, "IQG Control With An H_∞ Performance Bound: A Riccati Equation Approach," IEEE Trans. Autom. Contr., Vol. 34, 1989, to appear.
- 118. E. G. Collins, Jr., and D. C. Hyland, "Improved Robust Performance Bounds in Covariance Majorant Analysis," Int. J. Contr., 1989, to appear.

Submitted

119. D. S. Bernstein and W. M. Haddad, "Robust Stability and Performance for Fixed-Order Dynamic Compensation via the Optimal Projection Equations with Guaranteed Cost Bounds," submitted to Math. Contr. Sig. Sys.
120. D. C. Hyland and E. G. Collins, Jr., "An M-Matrix and Majorant Approach to Robust Stability and Performance Analysis for Systems with Structured Uncertainty," submitted to IEEE Trans. Autom. Contr.
121. D. S. Bernstein and W. M. Haddad, "Robust Decentralized Static Output Feedback," submitted to Sys. Contr. Lett.
122. D. S. Bernstein and I. G. Rosen, "Finite-Dimensional Approximation for Optimal Fixed-Order Compensation of Distributed Parameter Systems," submitted to Opt. Contr. Appl. Meth.
123. D. S. Bernstein and W. M. Haddad, "Robust Stability and Performance Analysis for State Space Systems via Quadratic Lyapunov Bounds," submitted to SIAM J. Matrix Anal. Appl.
124. D. C. Hyland and E. G. Collins, Jr., "A Majorant Approach to Robust Stability and Time-Domain Performance for Discrete-Time Systems," submitted to Math. Contr. Sig. Sys.
125. D. S. Bernstein and W. M. Haddad, "Optimal Reduced-Order State Estimation for Unstable Plants," submitted to Int. J. Contr.
126. W. M. Haddad and D. S. Bernstein, "Combined L_2/H_∞ Model Reduction," Amer. Contr. Conf., Pittsburgh, PA, June 1989.
127. D. S. Bernstein and W. M. Haddad, "Steady-State Kalman Filtering with an H_∞ Error Bound," Amer. Contr. Conf., Pittsburgh, PA, June 1989.
128. D. S. Bernstein and C. V. Hollot, "Robust Stability for Sampled-Data Control Systems," Amer. Contr. Conf., Pittsburgh, PA, June 1989.
129. W. M. Haddad and D. S. Bernstein, "Generalized Riccati Equations for H_∞ Controller Design with Structured Real-Valued Parameter Uncertainty," (abstract), Amer. Contr. Conf., Pittsburgh, PA, June 1989.
130. Y. Halevi, W. M. Haddad, and D. S. Bernstein, "A Riccati Equation Approach to the Singular LQG Problem," (abstract), Amer. Contr. Conf., Pittsburgh, PA, June 1989.
131. D. S. Bernstein, W. M. Haddad, and C. N. Nett, "Optimal Constrained-Structure Control Laws: A Basis for Integrated Design/Implementation of Nonlinear Digital Controllers," (abstract), Amer. Contr. Conf., Pittsburgh, PA, June 1989.
132. D. C. Hyland and E. G. Collins, Jr., "A Majorant Approach to Robust Stability and Time-Domain Performance for Discrete-Time Systems," submitted to Amer. Contr. Conf., Pittsburgh, PA, June 1989.

In Preparation

133. W. M. Haddad and D. S. Bernstein, "Optimal Reduced-Order Observer-Estimator Design With a Frequency-Domain Error Bound," in preparation for Advances in Control and Dynamic Systems, Vol. 31, Academic Press, 1989.
134. W. M. Haddad and D. S. Bernstein, "Optimal Reduced-Order Observer-Estimators," in preparation for AIAA J. Guid Contr. Dyn. Contr.
135. D. S. Bernstein and C. V. Hollot, "Robust Stability for Sampled-Data Control Systems," in preparation for IEEE Trans. Autom. Contr.
136. W. M. Haddad and D. S. Bernstein, "Robust LQG/ H_∞ Control Design with Structured Real-Valued Parameter Variations," in preparation for Int. J. Contr.
137. W. M. Haddad and D. S. Bernstein, "Robust Stability and Performance of Linear Dynamic Systems via Quadratic Lyapunov Bounds: Nonstrictly Proper Controller Synthesis," in preparation for Int. J. Contr.
138. Y. Halevi, W. M. Haddad, and D. S. Bernstein, "A Riccati Equation Approach to the Singular LQG Problem," in preparation for Automatica.
139. Y. Halevi, D. S. Bernstein, and W. M. Haddad, "Reduced-Order Observer-Estimators for Systems with Singular Measurement Noise," in preparation for IEEE Trans. Autom. Contr.
140. D. S. Bernstein and W. M. Haddad, Linear-Quadratic Feedback Control: Analysis and Design, in preparation for Lecture Notes in Control and Information Sciences, Springer-Verlag.

SECTION 7.0
PROGRAM PERSONNEL

7.0 Program Personnel

7.1 Dr. Dennis S. Bernstein

The principal investigator for this project was Dr. D. S. Bernstein. Dr. Bernstein received the Sc.B. in applied mathematics from Brown University and both the M.S. and Ph.D. degrees from the University of Michigan in control engineering. From 1982 to 1984 he was a staff member at Lincoln Laboratory, MIT, before joining Harris in 1984. He is a member of IEEE and SIAM.

7.2 Professor Wassim M. Haddad

Throughout this project Dr. Bernstein was assisted by Professor W. M. Haddad. Professor Haddad received the Ph.D. degree in 1987 from Florida Institute of Technology in mechanical engineering with specialization in systems and control. His doctoral research on robust control theory contributed directly to many of the developments achieved under this project. He is currently a faculty member at FIT and is a member of IEEE.

7.3 Acknowledgements

Many persons contributed either directly or indirectly to the results obtained under this project. At Harris we specifically acknowledge numerous helpful conversations with Dr. D. Hyland, Dr. J. Shipley, Dr. E. Collins, Mr. S. Richter, Mr. S. Greeley, Mr. L. Davis, Mr. A. Daubendiek, Mr. D. Phillips, and Mr. A. Tellez. We are also grateful to Mssrs. Richter, Greeley, Daubendiek, and Tellez for performing the numerical calculations in support of the project.

Finally, we are particularly indebted to Ms. J. Straehla who, without assistance, mastered PC TEX and singlehandedly produced every paper emanating from this project. Her contribution was exceptional and invaluable to the project. We also thank Ms. L. Ford for producing this final report.

APPENDIX A: OPUS Review Paper

88. D. S. Bernstein and D. C. Hyland, "Optimal Projection for Uncertain Systems (OPUS): A Unified Theory of Reduced-Order, Robust Control Design," in Large Space Structures: Dynamics and Control, S. N. Atluri and A. K. Amos, Eds., pp. 263-302, Springer-Verlag, New York, 1988.

Optimal Projection for Uncertain Systems (OPUS): A Unified Theory of Reduced-Order, Robust Control Design

Dennis S. Bernstein and David C. Hyland
Harris Corporation
Government Aerospace Systems Division
Melbourne, Florida 32902

This research was supported in part by the Air Force Office of Scientific Research under contracts F49620-86-C-0002 and F49620-86-C-0038.

Abstract

OPUS (Optimal Projection for Uncertain Systems) provides new machinery for designing active controllers for suppressing vibration in flexible structures. The purpose of this paper is to review this machinery and demonstrate its practical value in addressing the structural control problem.

1. Introduction

For many years it has been widely recognized that the desire to orbit large, lightweight space structures possessing high-performance capabilities would require active feedback control techniques. More generally, the need for such techniques may arise due to the combinations of either 1) moderate performance requirements for highly flexible structures with low-frequency modes or 2) stringent performance requirements for semi-rigid structures with relatively high-frequency modes (Figure 1). Applications include pointing, slewing, and aperture shape control for optical and RF systems.

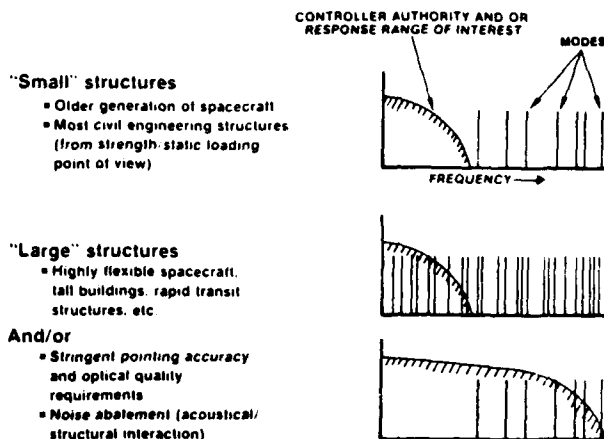


Figure 1. The Need for Active Structural Control Arises From Stringent Performance Requirements or Low-Frequency Modes

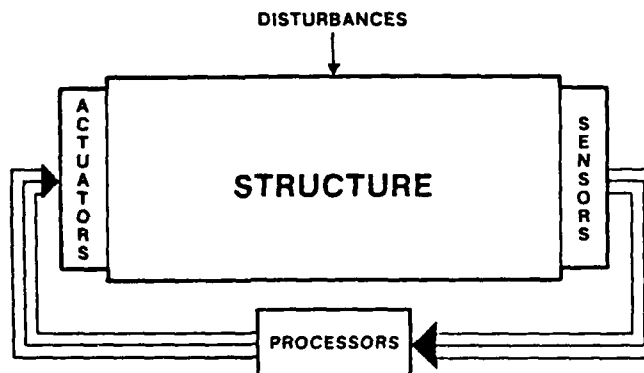


Figure 2. Vibration Control Systems Utilize Sensors, Processors and Actuators to Suppress Disturbances

The problem of active vibration suppression (Figure 2) entails the following considerations:

1. Multiple, highly coupled feedback loops. The potentially large number of sensors and actuators leads to a fully coupled multi-input, multi-output feedback control system.
2. Limited actuator power. The control authority available from on-board actuators is limited by weight, size, cost and power considerations.
3. High-dimensional models. Large structures subjected to broadband disturbances are typically represented by high-order finite element models.
4. Limited processor capacity. Reliability and cost considerations limit the processor capacity available for on-board real-time implementation of the control system.
5. Highly uncertain models with structured uncertainty. Finite element models often exhibit significant error particularly as modal frequency increases. Although modal testing and related identification methods may be used to improve modeling accuracy, residual uncertainty always remains and unpredictable on-orbit changes due to aging, thermal effects, etc., must be tolerated.

6. Stringent performance requirements. Since active space structure control is most relevant in precision applications, it can readily be expected that performance specifications will be particularly stringent.
7. Design efficiency. Because of implementation complexity due to the presence of multiple loops, high dimension, and high levels of uncertainty, the control design approach should efficiently utilize both synthesis and analysis techniques (Figure 3).

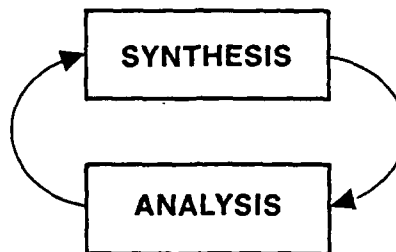
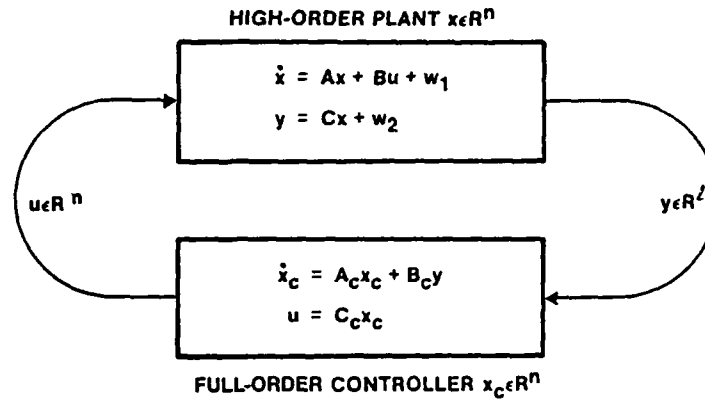


Figure 3. Control-System Design Must Efficiently Utilize Both Synthesis and Analysis Techniques

These considerations pose a considerable challenge to the state-of-the-art in control-design methodologies. For example, the presence of multiple, coupled feedback paths essentially precludes the effectiveness of single-loop design techniques. The sheer number of loops, their interaction, and the need to address a host of other issues render such methods inefficient and unwieldy.

In addition to the presence of multiple loops, the high dimensionality of dynamic models places a severe burden on control-design methodologies. For example, although LQG (linear-quadratic-Gaussian) design is applicable to multi-loop problems, such controllers are of the same order as the structural model (Figures 4 and 5). Thus LQG and similar high-order controllers can be expected to place an unacceptable computational burden on the real-time processing capability. Hence it is not surprising that a variety of techniques have been proposed to reduce the order of LQG controllers. A comparison of several such methods is given in [1].

All of the above difficulties are severely exacerbated by the fact that the dynamic (i.e., finite element) model upon which the control design is predicated may be highly inaccurate in spite of extensive modal identification. Hence, applicable control-design methodologies must account for modeling uncertainties by providing robust (i.e., insensitive) controllers. Furthermore, because of stringent



STEADY-STATE PERFORMANCE CRITERION

$$J(A_c, B_c, C_c) = \lim_{t \rightarrow \infty} E[x^T R_1 x + u^T R_2 u]$$

Figure 4. LQG Theory Addresses the Problem of Designing a Quadratically Optimal, Full-Order Dynamic Compensator

FULL-ORDER CONTROLLER GAINS

$$A_c = A - Q\bar{\Sigma} - \Sigma P$$

$$B_c = Q C^T V_2^{-1}$$

$$C_c = -R_2^{-1} B^T P$$

SEPARATED RICCATI EQUATIONS

$$0 = A Q + Q A^T + V_1 - Q \bar{\Sigma} Q \quad (\text{Kalman Filter})$$

$$0 = A^T P + P A + R_1 - P \Sigma P \quad (\text{Regulator})$$

$$\Sigma \equiv B R_2^{-1} B^T \quad \bar{\Sigma} \equiv C^T V_2^{-1} C$$

Figure 5. The Optimal Full-Order (LQG) Controller Is Determined by a Pair of Separated Riccati Equations

performance requirements, robust control design must avoid conservatism with respect to modeling uncertainty which may unnecessarily degrade performance. A salient example of conservatism is illustrated in Figure 6. If uncertainty in the modal frequency is complexified in a transfer function setting, then the resulting pole location uncertainty has the form of a disk. This disk, however, intersects the right half plane in violation of energy dissipation. Hence one source of conservatism is the inability to differentiate between physically distinct parameters such as modal frequency and modal damping.

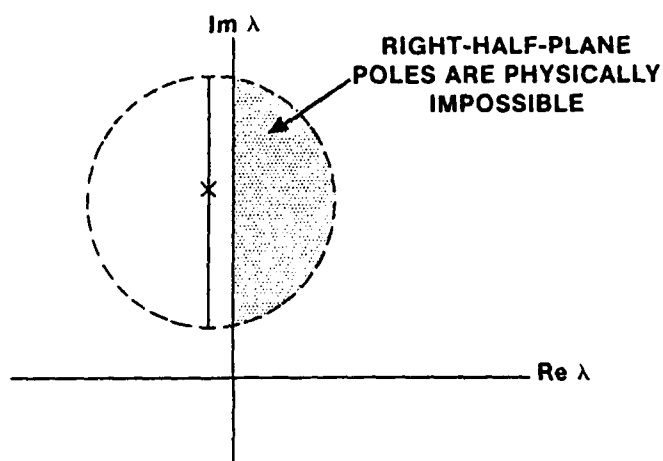


Figure 6. Complexification of Real Parameters May Lead to Robustness Conservatism

Although classical methods are inappropriate for vibration control, a wide variety of modern techniques are available. These include both multi-loop frequency-domain methods and time-domain techniques. A comprehensive review of such methods will not be attempted here. Rather, we shall merely point out aspects of several methods which motivate the philosophy of OPUS development.

As is well known, dynamic models can be transformed (at least in theory) between the frequency and time domains. Significant differences arise, however, in attempting to represent modeling errors. Specifically, model-error characterization of a particular type, which is natural and tractable in one domain, may become extremely cumbersome when transformed into the other domain. For example, consider a state space model with parameter uncertainties arising in the system matrices (A, B, C) . Upon transforming to a frequency domain model $G(s) = C(sI - A)^{-1}B$ the parametric uncertainties may perturb the transfer function coefficients in a

complicated manner. A more natural measure of uncertainty for transfer functions has been developed in [2] where system uncertainty in the frequency domain is modeled by means of normed neighborhoods in the H-infinity topology. There are limitations with this approach, however, in designing controllers for vibration suppression. For example, as shown in Figure 6, complexification of real-parameter uncertainties such as modal frequencies may yield unnecessary conservatism, while norm bounds often fail to preserve the physical structure of parameter variations. A case in point is the lightly damped oscillator. As shown in [A42], norm bounds predict stability over a frequency range on the order of the damping while in fact the oscillator is unconditionally stable. Furthermore, with regard to processor throughput tradeoffs, modern frequency-domain methods typically yield high-order controllers.

Although LQG addresses performance/actuator and performance/sensor tradeoffs in a multi-loop setting, it fails to incorporate modeling uncertainty. Thus it is not surprising, as shown in [3], that LQG designs fail to possess guaranteed gain margin. Since LQG designs lack such margins, attempts have been made to apply frequency-domain techniques to improve their characteristics. One such method, known as LQG/LTR ([4,5]) seeks to recover the gain margin of full-state-feedback controllers. Specifically, full-state-feedback LQR controllers are guaranteed to remain stable in the face of perturbations of the input matrix B of the form αB where $\alpha \in [1/2, \infty)$. As shown in [6,7], however, the full-state-feedback gain margin fails to provide robustness with respect to perturbations which are not of this form. For instance, the example given in [6] with $B = [0 \ 1]^T$ can be destabilized for suitable performance weightings with perturbation $B(\epsilon) = [\epsilon \ 1]^T$ for arbitrarily small ϵ in spite of the 6 dB margin. Furthermore, since LQG/LTR loop shaping is based upon singular value norm bounds, treatment of physically meaningful real parameter variations may lead to unnecessary conservatism. Several approaches have been proposed for circumventing these difficulties (see, e.g., [8]).

The importance of addressing the problem of structured uncertainty in finite element models cannot be overemphasized. Structural characteristics such as modal frequencies, damping ratios, and mode shapes appear explicitly in (A,B,C) state-space models as physically meaningful parameters. Uncertainty in mode shapes, for example, which appear as columns of the B matrix, cannot in general be expected to be of a multiplicative form in accordance with traditional gain-margin specifications. This is precisely the problem illustrated by the example of [6] discussed above. Furthermore, uncertainties in modal frequencies and damping ratios must be carefully differentiated since, roughly speaking, modal frequency uncertainties affect only the imaginary part of the pole location while damping uncertainty affects the real part. Although these and related observations

concerning uncertainty in the dynamic characteristics of lightly damped structures may be self evident, they have remained largely unexploited in standard control-design methods.

2. OPUS: New Machinery for Control-System Design

In view of the ability of LQG theory to synthesize dynamic controllers for multi-input, multi-output controllers, it is not surprising that LQG forms the basis for a variety of structural control methods. However, as discussed previously, LQG lacks the ability to address performance/processor and performance/robustness tradeoffs. This situation has thus motivated the development of numerous variants of LQG which entail additional procedures which attempt to remedy these defects. OPUS, however, is distinctly different. Rather than append additional procedures to LQG design, OPUS extends LQG theory itself by generalizing the basic underlying machinery.

As shown in Figure 5, the basic machinery of LQG consists of a pair of separated Riccati equations whose solutions serve to directly and explicitly synthesize the gains of an optimal dynamic compensator. The contribution of OPUS is to directly expand this machinery. The overall approach is illustrated in Figure 7 which portrays two distinct generalizations of the basic LQG machinery. As Figure 7 illustrates, these generalizations can be developed individually when either low-order or robust controllers are desired. The appealing aspect of OPUS, however, is the ability to extend LQG to address both problems simultaneously in a unified manner.

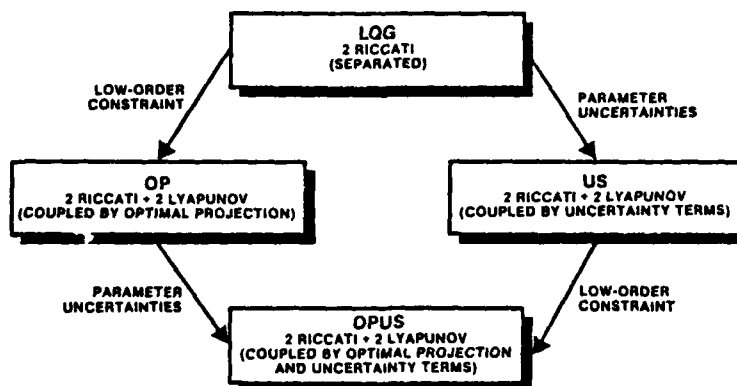


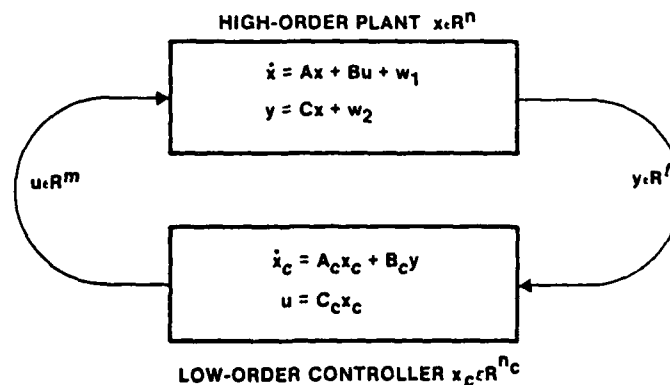
Figure 7. The Standard LQG Result Is Generalized by Both the Fixed-Order Constraint and Modeling of Parameter Uncertainties

In the following sections the generalizations depicted in Figure 7 will be reviewed following the left branch. That is, the optimal projection approach to reduced-order controller design will first be discussed in Section 3 without introducing plant uncertainties. In Section 4 the reduced-order constraint will be retained while considering, in addition, uncertainties in the system model. In each case the discussion will focus on the underlying ideas with a minimum of technical detail.

Clearly, in order for a novel design methodology to be of practical value it must be computationally tractable. Hence Section 5 will present an overview of the current state of algorithm development for solving the OPUS design equations. Finally, Section 6 will briefly summarize further OPUS generalizations of LQG theory which are relevant to structural control.

3. Extensions of LQG to Reduced-Order Dynamic Compensation

The simplest, most direct way to obtain optimal reduced-order controllers is to redevelop the standard LQG result in the presence of a constraint on controller dimension (Figure 8). The mathematical technique required to do this is remarkably straightforward. Specifically, the structure and order of the controller are fixed and the performance is optimized with respect to the controller gains. The resulting necessary conditions obtained using Lagrange multipliers thus characterize the optimal gains.



STEADY-STATE PERFORMANCE CRITERION

$$J(A_c, B_c, C_c) = \lim_{t \rightarrow \infty} E[x^T R_1 x + u^T R_2 u]$$

Figure 8. In Accordance With On-Board Processor Requirements, a Reduced-Order Constraint Is Imposed on the Dimension of the Dynamic Compensator

This parameter optimization approach as such is not new and was investigated extensively in the 1970's. Typically, however, the optimality conditions were found to be complex and unwieldy while offering little insight and requiring gradient search methods for numerical solution.

One curious aspect of the parameter optimization literature is that no attempt was made to actually use this direct method to rederive the LQG result itself. Such an exercise, it may be surmised, might reveal hidden structure within the optimality conditions which would shed light on the reduced-order case. Indeed, such an approach led to the realization that an oblique projection (idempotent matrix) is the key to unlocking the unwieldy optimality conditions ([A7,A17]). Although the result is mathematically straightforward, it is by no means obvious since in the full-order (LQG) case the projection is the identity and hence not readily apparent.

By exploiting the presence of the projection, the necessary conditions can be transformed into a coupled system of four algebraic matrix equations consisting of a pair of modified Riccati equations and a pair of modified Lyapunov equations (Figure 9). The coupling is via the oblique projection τ which appears in all four equations and which is determined by the solutions \hat{Q} and \hat{P} of the modified Lyapunov equations. A satisfying feature of the optimality conditions is that in the full-order case the projection becomes the identity, the modified Lyapunov equations drop out, and, since $\tau_1 = 0$, the modified Riccati equations specialize to the usual separated Riccati equations of LQG theory. Since, furthermore, $G = \Gamma = nxn$ identity, the standard LQG gain expressions are recovered.

Although the modified Riccati equations specialize to the standard Riccati equations in the full-order case, the modified Lyapunov equations have no counterpart in the standard theory. The role of these equations can be understood by considering the problem of optimal model reduction alone. For this problem the optimal reduced-order model is characterized by a pair of coupled modified Lyapunov equations (see [A22]). Thus the modified Lyapunov equations arising in the reduced-order dynamic-compensation problem are directly analogous to the modified Lyapunov equations arising in model reduction alone. The modified Lyapunov equations arising in the control problem, however, are intimately coupled with the modified Riccati equations. Hence it cannot be expected that reduced-order control-design techniques based upon LQG will generally yield optimal fixed-order controllers (Figure 10). It is interesting to note that several such methods discussed in [1] are based upon balancing which was shown in [A22] to be suboptimal with respect to the quadratic (least squares) optimality criterion.

REDUCED-ORDER CONTROLLER GAINS

$$A_c = I'(A - Q\bar{\Sigma} - \Sigma P)G^T$$

$$B_c = I'QC^TV_2^{-1}$$

$$C_c = -R_2^{-1}B^TPG^T$$

COUPLED RICCATI/LYAPUNOV EQUATIONS

$$0 = AQ + QA^T + V_1 - Q\bar{\Sigma}Q + r_1Q\bar{\Sigma}Qr_1^T$$

$$0 = ATP + PA + R_1 - P\Sigma P + r_1^TP\Sigma Pr_1$$

$$0 = (A - \Sigma P)\hat{Q} + \hat{Q}(A - \Sigma P)^T + Q\bar{\Sigma}Q - r_1Q\bar{\Sigma}Qr_1^T$$

$$0 = (A - Q\bar{\Sigma})^T\hat{P} + \hat{P}(A - Q\bar{\Sigma}) + P\Sigma P - r_1^TP\Sigma Pr_1$$

$$\text{rank } \hat{Q} = \text{rank } \hat{P} = \text{rank } \hat{Q}\hat{P} = n_c$$

$$\hat{Q}\hat{P} = G^TM I' \quad I'G^T = I_{n_c}$$

$$r = GTI' = \hat{Q}\hat{P}(\hat{Q}\hat{P})^\# \quad r_1 = I_n - r$$

$$\Sigma = BR_2^{-1}B^T \quad \bar{\Sigma} = C^TV_2^{-1}C$$

Figure 9. The Optimal Reduced-Order Compensator Is Determined by a Pair of Modified Riccati Equations and a Pair of Modified Lyapunov Equations Coupled by the Oblique Projection r

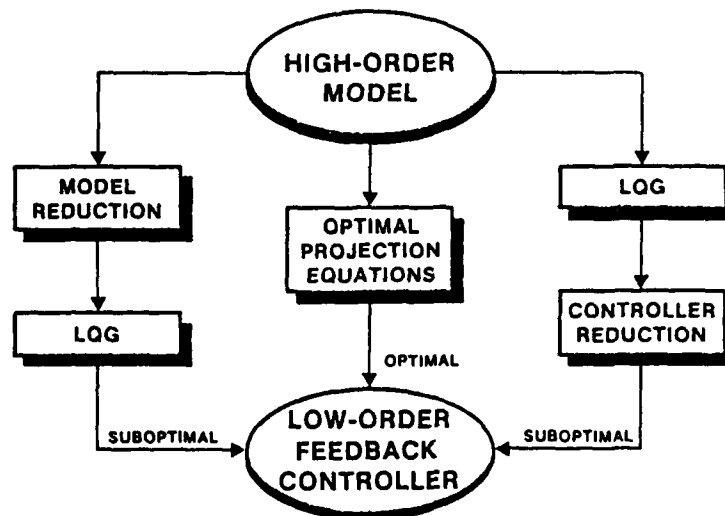


Figure 10. The Optimal Projection Equations Provide a Direct Path to Optimal Reduced-Order Dynamic Compensators

In summary, the optimal projection equations for reduced-order dynamic compensation comprise a direct extension of the basic LQG machinery to the reduced-order control problem. The design equations, which reduce to the standard LQG result in the full-order case, provide direct synthesis of optimal reduced-order controllers in accordance with implementation constraints.

4. Extensions of LQG to Uncertain Modeling

Two fundamental sources of error in modeling flexible structures are truncated modes and parameter uncertainties. Since the optimal projection approach permits the utilization of the full dynamics model, modal truncation can be largely avoided. There remains, however, a tendency to truncate poorly known modes and thus it is essential to incorporate a model of parameter uncertainty in both well-known and poorly known components of the system. Hence the problem formulation of Figure 8 is now generalized in Figure 11 to include uncertain parameters σ_j appearing in the A, B and C matrices. The parameter σ_i is assumed to lie within the interval $[-\delta_i, \delta_i]$ in accordance with identification accuracy. Clearly, when uncertainty is absent, i.e., when $A_i, B_i, C_i = 0$, the reduced-order design problem of Figure 8 is recovered.

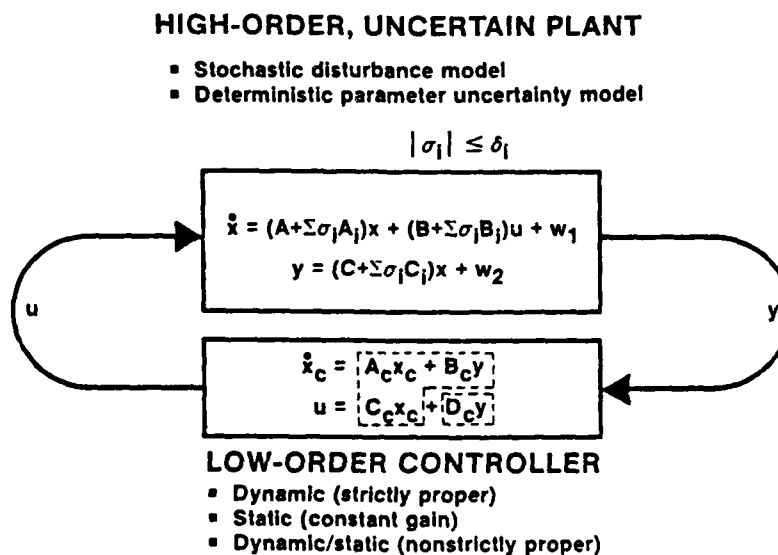
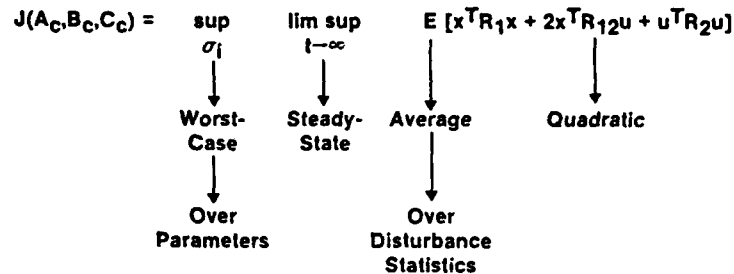


Figure 11. Robust Optimal Projection Design Is Based Upon a Hybrid Uncertainty Model Involving a Deterministic Parameter Uncertainty Model and a Stochastic Disturbance Model

A salient feature of the design model is that uncertainty is modeled in two distinctly different ways. External uncertainty appearing as additive white noise is modeled stochastically. Such a model appears appropriate for disturbances such as coolant flow for which only power spectral data are available. On the other hand, internal uncertainty appearing as parameter variations is modeled deterministically. Such a model appears appropriate for uncertainty arising from directly measurable quantities such as mass and stiffness. Thus the overall uncertainty model is hybrid in the sense that it utilizes both deterministic and stochastic characterizations of uncertainty.

A natural performance measure which accounts for both types of uncertainty characterization involves the usual LQG quadratic criterion averaged over the disturbance statistics and then maximized over the uncertain parameters (Figure 12). Hence this performance measure incorporates on the average and worst case aspects in accordance with physical considerations.

PERFORMANCE CRITERION



ROBUST PERFORMANCE PROBLEM

Minimize $J(A_c, B_c, C_c)$ over the class of robustly stabilizing controllers (A_c, B_c, C_c)

Figure 12. Performance Is Defined To Be Worst Case Over the Uncertain Parameters and Average Over the Disturbance Statistics

The resulting Robust Performance Problem thus involves determining the gains (A_c, B_c, C_c) to minimize the performance J . The static gain D_c can also be included but will not be discussed here. Despite the apparent complexity of the problem, remarkably simple techniques can be used. Specifically, first note that after taking the expected value the performance J has the form

$$J(A_c, B_c, C_c) = \sup_{\sigma_i} \limsup_{t \rightarrow \infty} \text{tr } \bar{Q}(t) \bar{R}, \quad (4.1)$$

where "tr" denotes trace of a matrix, $\bar{Q}(t)$ is the covariance of the closed-loop system, and \bar{R} is an augmented weighting matrix composed of R_1 , R_{12} and R_2 . The covariance $\bar{Q}(t)$ satisfies the standard Lyapunov differential equation

$$\dot{\bar{Q}} = (\bar{A} + \sum \sigma_i \bar{A}_i) \bar{Q} + \bar{Q} (\bar{A} + \sum \sigma_i \bar{A}_i)^T + \bar{V}, \quad (4.2)$$

where \bar{A} is the closed-loop dynamics, \bar{A}_i is composed of A_i , B_i and C_i , and \bar{V} is the intensity of external disturbances for the closed-loop system including the plant and measurement noise.

Two distinct approaches to this problem will be considered. The first involves bounding the performance over the class of parameter uncertainties and then choosing the gains to minimize the bound. Since bounding precedes control design this approach is known as robust design via a priori performance bounds. The second approach involves exploiting the nondestabilizing nature of structural systems via weak subsystem interaction.

4.1 Robust Design Via A Priori Performance Bounds

The key step in bounding the performance (4.1) is to replace (4.2) by a modified Lyapunov differential equation of the form

$$\dot{\bar{Q}} = \bar{A}\bar{Q} + \bar{Q}\bar{A}^T + \Psi(\bar{Q}) + \bar{V}, \quad (4.3)$$

where the bound Ψ satisfies the inequality

$$\sum \sigma_i (\bar{A}_i \bar{Q} + \bar{Q} \bar{A}_i^T) \leq \Psi(\bar{Q}) \quad (4.4)$$

over the range of uncertain parameters σ_i and for all candidate feedback gains. Note that the inequality (4.4) is defined in the sense of nonnegative-definite matrices. Now rewrite (4.3) by appropriate addition and subtraction as

$$\dot{\bar{Q}} = (\bar{A} + \sum \sigma_i \bar{A}_i) \bar{Q} + \bar{Q} (\bar{A} + \sum \sigma_i \bar{A}_i)^T + \Psi(\bar{Q}) - \sum \sigma_i (\bar{A}_i \bar{Q} + \bar{Q} \bar{A}_i^T) + \bar{V}. \quad (4.5)$$

Now subtract (4.2) from (4.5) to obtain

$$\dot{\bar{Q}} - \dot{\bar{Q}} = (\bar{A} + \sum \sigma_i \bar{A}_i) (\bar{Q} - \bar{Q}) + (\bar{Q} - \bar{Q}) (\bar{A} + \sum \sigma_i \bar{A}_i)^T + \Psi(\bar{Q}) - \sum \sigma_i (\bar{A}_i \bar{Q} + \bar{Q} \bar{A}_i^T). \quad (4.6)$$

Since by (4.4) the term

$$\Psi(\bar{Q}) = \sum \sigma_i (\bar{A}_i \bar{Q} + \bar{Q} \bar{A}_i^T) \quad (4.7)$$

is nonnegative definite, it follows immediately that

$$\bar{Q} \leq \bar{Q} \quad (4.8)$$

over the class of uncertain parameters. Thus the performance (4.1) can be bounded by

$$J(A_c, B_c, C_c) \leq \underline{J}(A_c, B_c, C_c) \triangleq \lim_{t \rightarrow \infty} \text{tr } \bar{Q} R. \quad (4.9)$$

The auxiliary cost \underline{J} is thus guaranteed to bound the actual cost J . This leads to the Auxiliary Minimization Problem: Minimize the auxiliary cost \underline{J} over the controller gains. The advantage of this approach is that necessary conditions for the Auxiliary Minimization Problem effectively serve as sufficient conditions for robust performance in the original problem. Since the bounding step precedes the optimization procedure, this approach is referred to as robust design via a priori performance bounds. This procedure is philosophically similar to guaranteed cost control ([9,10]). Note that since bounding precedes optimization, the bound (4.4) must hold for all gains since the optimal gains are yet to be determined.

To obtain sufficient conditions for robust stability, the bounding function Ψ must be specified. Since the ordering of nonnegative-definite matrices appearing in (4.4) is not a total ordering, a unique lowest bound should not be expected. Furthermore, each differentiable bound leads to a fundamental extension of the optimal projection equations and thus of the basic LQG machinery. In work thus far, two bounds have been extensively investigated. Only one bound, the right shift/multiplicative white noise bound, will be discussed here. The structured stability radius bound introduced in [11,12] is discussed in [A43].

The right shift/multiplicative white noise bound investigated in [A29,A41] is given by

$$\Psi(\bar{Q}) = \sum \delta_j (\alpha_j \bar{Q} + \alpha_j^{-1} \bar{A}_j \bar{Q} \bar{A}_j^T), \quad (4.10)$$

where $\alpha_j > 0$ are arbitrary scalars. Note that this bound consists of two distinct parts which must appear in an appropriate ratio. The first term $\alpha_j \bar{Q}$ arises naturally when an exponential time weighting $e^{\alpha_j t}$ is included in the performance measure. As is well known ([13]) this leads to a prescribed uniform stability margin for the

closed-loop system (Figure 13). A uniform stability margin, no matter how large, however, does not guarantee robustness with respect to arbitrary parameter variations. The complementary second term $\alpha_i^{-1} \bar{A}_i \bar{Q} \bar{A}_i^T$ is crucial in this regard.

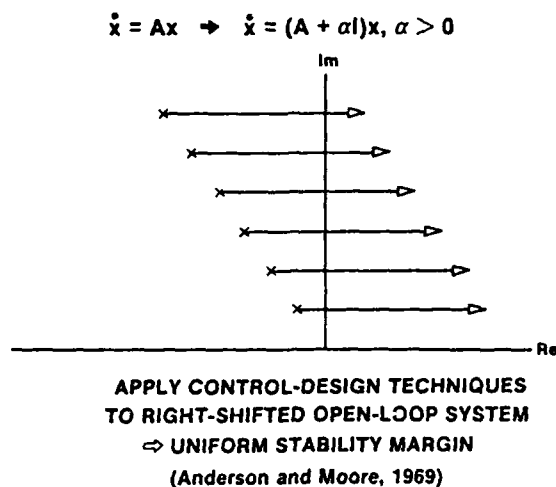


Figure 13. Open-Loop Right-Shifted Dynamics Arising From Exponential Cost Weighting Lead to a Uniform Closed-Loop Stability Margin

Although terms of the form $\bar{A}_i \bar{Q} \bar{A}_i^T$ are unfamiliar in robust control design, they arise naturally in stochastic differential equations with multiplicative white noise. That is, if the uncertain parameters σ_i are replaced by white noise processes entering multiplicatively rather than additively, then the covariance equation for \bar{Q} automatically includes terms of the form $\bar{A}_i \bar{Q} \bar{A}_i^T$. The literature on systems with multiplicative white noise is quite extensive; see [A38] for references. It should be stressed, however, that for our purposes the multiplicative white noise model is not interpreted literally as having physical significance. Rather, multiplicative white noise can be thought of as a useful design model which correctly captures the impact of uncertainty on the performance functional via the state covariance. Furthermore, just as the right shift term alone does not guarantee robustness, neither does the multiplicative white noise term. Both terms must appear simultaneously. Roughly speaking, since multiplicative white noise disturbs the plant through uncertain parameters, the closed-loop system is automatically desensitized to actual parameter variations.

After incorporating the right shift/multiplicative white noise bound (4.10) into (4.3) to obtain a bound \underline{J} for the performance, the optimal projection equations can be rederived following exactly the same parameter optimization procedure discussed in Section 3. Again, the mathematics required is but a straightforward application of Lagrange multipliers. The additional bounding terms are carried through the derivation to yield a direct generalization of the optimal projection equations shown in Figure 14 with gains given in Figure 15.

$$0 = A_s Q + Q A_s^T + \lambda Q A^T + V_1 + (A B R_{2s}^{-1} P_s) \hat{Q} (A - B R_{2s}^{-1} P_s)^T - Q_s V_{2s}^{-1} Q_s^T + r_1 Q_s V_{2s}^{-1} Q_s^T r_1^T$$

$$0 = A_s^T P + P A_s + \lambda^T P A + R_1 + (A - Q_s V_{2s}^{-1} C_s)^T \hat{P} (A - Q_s V_{2s}^{-1} C_s) - P_s^T R_{2s}^{-1} P_s + r_1^T P_s^T R_{2s}^{-1} P_s r_1$$

$$0 = (A_s - B_s R_{2s}^{-1} P_s) \hat{Q} + \hat{Q} (A_s - B_s R_{2s}^{-1} P_s)^T + Q_s V_{2s}^{-1} Q_s^T - r_1 Q_s V_{2s}^{-1} Q_s^T r_1^T$$

$$0 = (A_s - Q_s V_{2s}^{-1} C_s)^T \hat{P} + \hat{P} (A_s - Q_s V_{2s}^{-1} C_s) + P_s^T R_{2s}^{-1} P_s - r_1^T P_s^T R_{2s}^{-1} P_s r_1$$

Figure 14. The Robustified Optimal Projection Design Equations Account for Both Reduced-Order Dynamic Compensation and Parametric Uncertainty

GAINS

$$A_c = P(A_s - B_s R_{2s}^{-1} P_s - Q_s V_{2s}^{-1} C_s) G^T$$

$$B_c = P Q_s V_{2s}^{-1}$$

$$C_c = -R_{2s}^{-1} P_s G^T$$

NOTATION

$$\hat{\hat{Q}} P = G^T M \Gamma, \quad \Gamma G^T = I_{n_c} \quad (\Rightarrow r = G^T \Gamma = r^2)$$

$$\lambda Q A^T = \sum_{i=1}^p A_i Q A_i^T, \quad \lambda Q B = \sum_{i=1}^p A_i Q B_i, \text{ etc.}$$

$$R_{2s} = R_2 + B^T (P + \hat{P}) B \quad V_{2s} = V_2 + C(Q + \hat{Q}) C^T$$

$$Q_s = Q C_s^T + V_{12} + \lambda(Q + \hat{Q}) C^T \quad P_s = B_s^T P + R_{12} + B^T (P + \hat{P}) A$$

Figure 15. The OPUS Controller Gains Are Explicitly Characterized as a Direct Generalization of the Classical LQG Gains

The robustified optimal projection equations comprise a system of four matrix equations coupled by both the optimal projection and uncertainty terms. When the uncertainty terms are absent, the optimal projection equations of Figure 9 are immediately recovered. On the other hand, if the order of the controller is set equal to the order of the plant, then all terms involving τ_1 can be deleted. However, in this case the modified Lyapunov equations do not drop out since \hat{Q} and \hat{P} still appear in the modified Riccati equations. Hence the basic machinery of LQC is again extended to include a pair of Lyapunov equations coupled to a generalization of the standard LQC equations. It is interesting to note that a related result in the context of multiplicative noise also appeared in the Soviet literature ([14]). It should also be pointed out that although the modified Lyapunov equations arising in the reduced-order control-design problem have analogues in model reduction, the modified Lyapunov equations appearing in the full-order robustified equations represent new machinery not anticipated in robustness theories. Hence using straightforward mathematical techniques, the basic LQC machinery has again been extended in novel directions.

Solving the design equations shown in Figures 14 and 15 yields controllers with guaranteed levels of robustness. The actual robustness levels may, however, be larger than specified by a priori bounds. Thus, to achieve desired robustification levels for the uncertainty structure specified by the a priori bounds, the design procedure may be utilized within an iterative synthesis/analysis procedure (Figure 16).

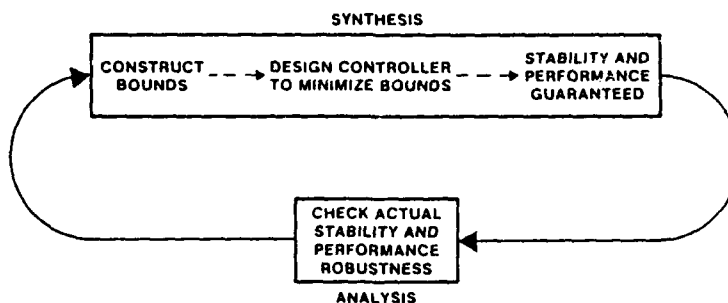


Figure 16. Optimal Projection/Guaranteed Cost Control Provides Direct Synthesis of Robust Dynamic Compensators

4.2 Robust Design Via Weak Subsystem Interaction

The mechanism by which LQC was robustified in Section 4.1 involved bounding the performance over the class of parameter uncertainties and then determining optimal controller gains for the performance bound. As discussed in Section 2,

however, flexible structures possess special properties which may, in addition, be exploited to achieve robustness. Specifically, aside from rigid-body modes, energy dissipation implies that mechanical structures are open-loop stable regardless of the level of uncertainty. That is, flexible structures possess only nondestabilizing uncertainties. Hence, in the closed loop, a given controller may or may not render a particular uncertainty destabilizing. A priori bounds on controller performance must, however, be valid for all gains since bounding precedes optimization. Hence, a priori bounding may in certain cases fail to exploit nondestabilizing uncertainties.

A familiar example of a nondestabilizing uncertainty involves uncertain modal frequencies. Such an uncertainty will not, of course, destabilize an uncontrolled (open-loop) structure. If particular modal frequencies are poorly known then it is clearly advisable to avoid applying high authority control. Hence, rather than the right-shift approach of Figure 13, it appears advantageous (although, at first, counterintuitive) to utilize just the opposite, namely, a left shift (Figure 17). Furthermore, in view of the fact that uncertainty usually increases with modal frequency (Figure 18), a variable left shift appears to be more appropriate than a uniform left shift. By left-shifting high-frequency poorly known modes, the control-system design procedure applies correspondingly reduced authority to modes "perceived" as highly damped. Hence the variable left shift can be roughly thought of as a device for achieving suitable authority rolloff. As will be seen, however, the underlying robustification mechanism, namely, weak subsystem interaction, is far more subtle than the approach of classical rolloff techniques. It is also interesting to note that the weak subsystem interaction approach to robustness is entirely distinct from classical robustness approaches which utilize high loop gain to reduce sensitivity.

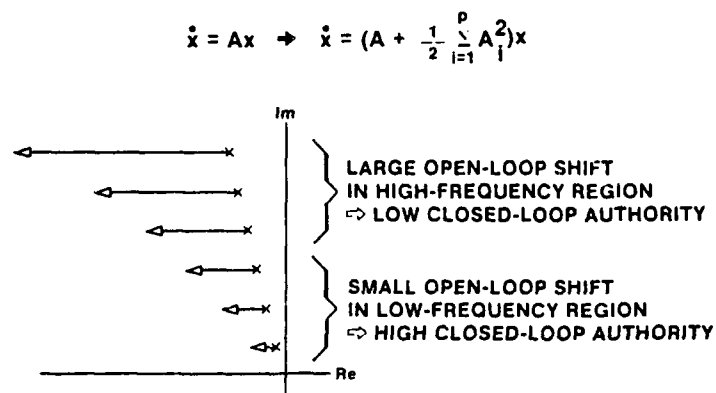


Figure 17. A Variable Left Shift Exploits Open-Loop Nondestabilizing Uncertainties

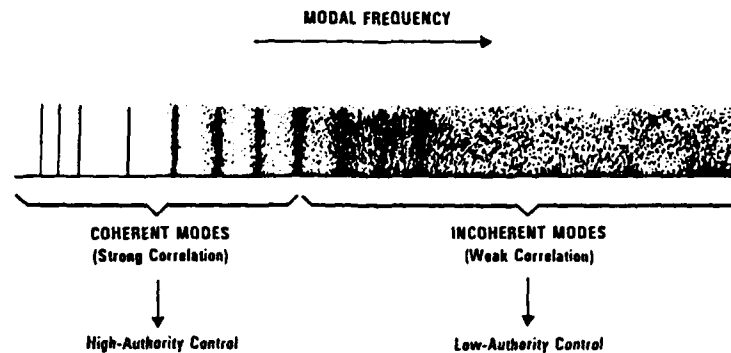


Figure 18. Modal Uncertainty Generally Increases With Frequency

A variable left shift can readily be introduced into the robustified optimal projection design equations by replacing A by

$$A_s = A + \frac{1}{2} \sum A_j^2 \quad (4.11)$$

where A_j denotes the structure of modal frequency uncertainty (Figure 19). Most interestingly, such a modification of the dynamics matrix arises naturally from a multiplicative white noise model defined not in the usual Ito sense but rather in the sense of Stratonovich. Thus, as in the a priori bounding approach, a stochastic

$$\begin{aligned}
 A &= \begin{bmatrix} -\eta_1 & \omega_1 & 1 & 0 \\ -\omega_1 & -\eta_1 & 1 & 0 \\ 0 & 0 & -\eta_2 & \omega_2 \\ & & -\omega_2 & -\eta_2 \end{bmatrix} & A_1 = \gamma_1 \begin{bmatrix} 0 & & & \\ & 0 & & \\ & & 0 & 1 \\ & & -1 & 0 \\ & & & 0 & 0 \end{bmatrix} \\
 \Rightarrow A_1^2 = \gamma_1^2 \begin{bmatrix} 0 & & & \\ & 0 & & \\ & & 0 & -1 \\ & & 1 & 0 \\ & & & 0 & 0 \end{bmatrix} & \Rightarrow \frac{1}{2} \sum A_1^2 = \begin{bmatrix} -\frac{1}{2} \gamma_1^2 \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} & 0 \\ 0 & -\frac{1}{2} \gamma_2^2 \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \end{bmatrix} \\
 \Rightarrow A + \frac{1}{2} \sum A_j^2 &\longleftrightarrow \text{Variable Left Shift}
 \end{aligned}$$

Figure 19. For Modal Systems With Frequency Uncertainty the Stratonovich Correction Corresponds to a Variable Left Shift

model serves to suggest a mechanism for robustification (Figure 20). Again it is important to stress that the multiplicative white noise model is not interpreted literally as having physical significance, but rather can be thought of as a useful design model which correctly captures the impact of uncertainty on the performance functional via the state covariance.

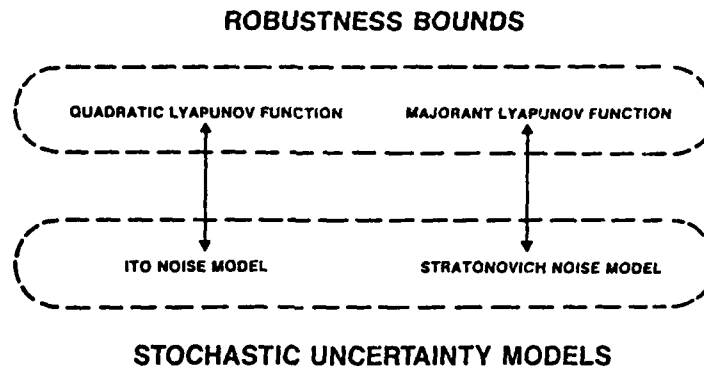


Figure 20. Stochastic Models and Robustness Bounds Are Fundamentally Related

In earlier work the Stratonovich dynamics model was justified by means of the minimum information/maximum entropy approach ([A1-A15]). A central result of the maximum entropy approach is that the high authority/low authority transition of a vibration control system from well-known low-frequency modes to poorly known high-frequency modes (Figure 18) is directly reflected in the structure of the state covariance matrix (Figure 21). A full-state feedback design applied to a simply

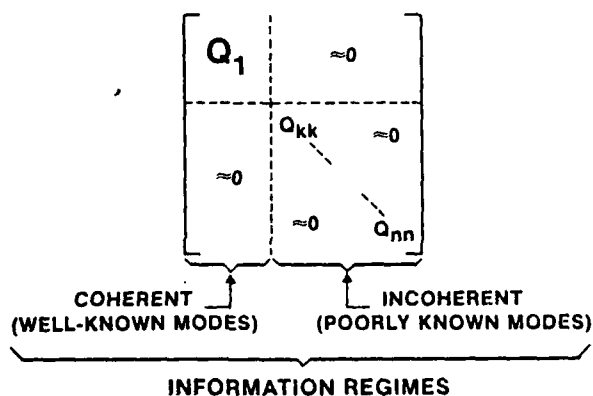


Figure 21. Frequency Uncertainties in the Stratonovich Model Lead to Suppressed Cross Correlation in the Steady-State Covariance

supported beam illustrates this point (Figure 22). By assuming that uncertainty in modal frequencies increases linearly with frequency, the structure of the covariance matrix leads directly to the control gains illustrated in Figure 23. Note that in the high-frequency region the position gains are essentially zero and thus the control law approaches positive-real energy dissipative rate feedback. This, of course, is precisely the type of structural controller expected in the presence of poor modeling information. Of course, any effective control-design theory for active vibration suppression in flexible structures should produce energy dissipative controllers when structural modeling information is highly uncertain.

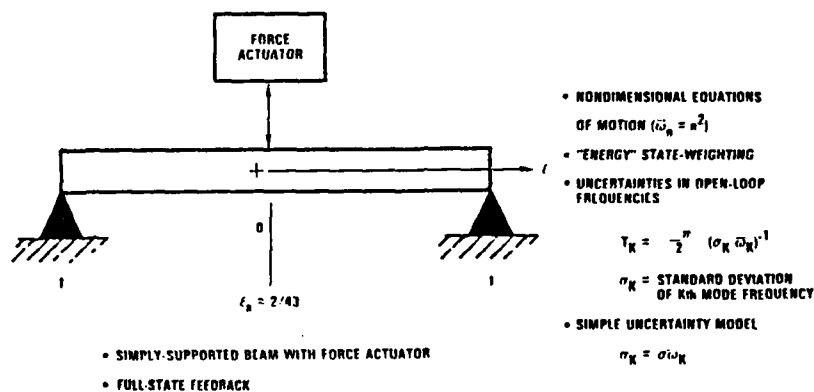


Figure 22. The Effects of Frequency Uncertainties Can Be Illustrated for a One-Dimensional Beam With Idealized Full-State Feedback

To carry out robustified optimal projection design in the presence of left-shifted open-loop dynamics, it is only necessary to utilize the left-shifted dynamics matrix (4.11) in place of the right-shifted matrix. All of the robustified optimal projection machinery, including gain expressions, can be utilized directly. It is also important to stress that the left shift must be used in conjunction with terms of the form $\bar{A}_i \bar{Q} \bar{A}_i^T$.

One explanation for the mechanism by which robustification is achieved is illustrated in Figure 24. By left shifting the open-loop dynamics within the design process, the compensator poles are similarly left-shifted. Thus the compensator poles are effectively moved further into the left half plane away from the actual plant poles. Since the interaction between compensator and plant poles is weakened, the closed-loop system is correspondingly robustified with respect to uncertainties in the plant pole locations. A sensitivity analysis of this mechanism utilizing a uniform left shift in the context of LQG design is given in [15].

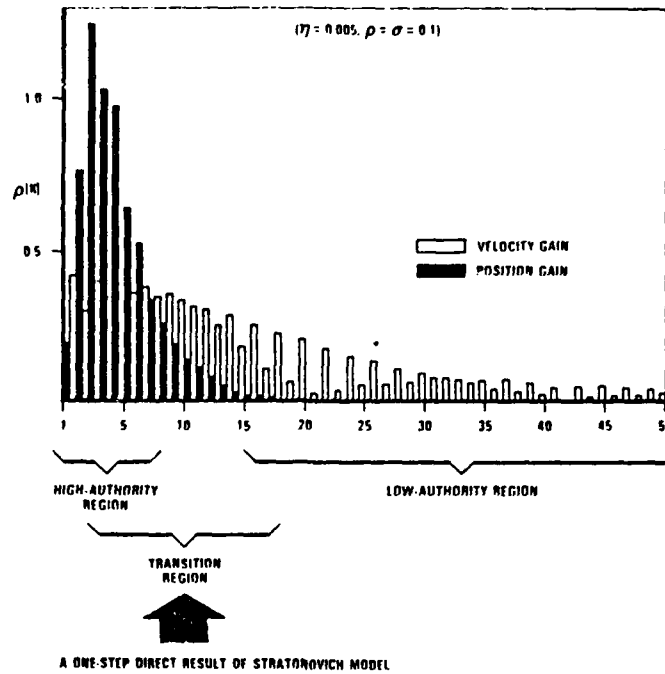


Figure 23. The Maximum Entropy Controller Approaches Rate Feedback in the Limit of Poor Modeling Information (High Uncertainty)

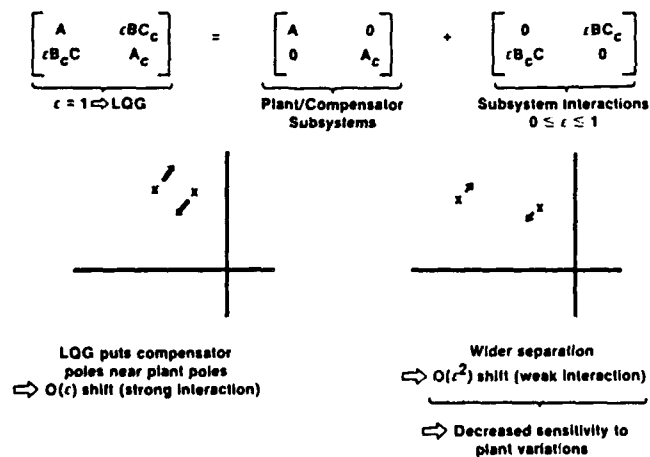


Figure 24. The Stratonovich Variable Left-Shift Model Effectively Places the Compensator Poles Further Into the Left Half Plane Where Plant/Compensator Interaction Is Weakened

As discussed above, the left-shift approach exploits open-loop nondestabilizing uncertainties and thus cannot operate through a priori bounding. Thus the actual level of robustification achieved from the robustified optimal projection equations for a given level of uncertainty modeling cannot be predicted a priori, i.e., in advance of control design. Indeed, this situation is to be expected when nondestabilizing uncertainties are exploited in a nonconservative design theory. Thus a suitable robust analysis technique is required for nonconservatively determining the robustification of the closed-loop system with respect to open-loop nondestabilizing uncertainties.

A suitable robustness analysis technique, known as majorant Lyapunov analysis, has indeed been developed ([A42]). Essentially, this technique employs a new type of Lyapunov function for assessing robustness due to weak subsystem interaction. The underlying machinery consists of the block-norm matrix which is a nonnegative matrix each of whose elements is the norm of a block of a suitably partitioned matrix (Figure 25). A matrix which bounds the block-norm matrix in the sense of nonnegative matrices, i.e., element by element, is known as a majorant. Majorants were introduced in [16] and were applied to stability analysis of integration algorithms for ODE's in [17].

(Ostrowski, 1961; Dahlquist, 1983)

$$M = \begin{bmatrix} M_1 & M_{12} & \cdots \\ M_{21} & M_2 & \ddots \\ \vdots & & \ddots \end{bmatrix}$$

$$m = \begin{bmatrix} \|M_1\| & \|M_{12}\| & \cdots \\ \|M_{21}\| & \|M_2\| & \ddots \\ \vdots & & \ddots \end{bmatrix}$$

NONNEGATIVE CONE ORDERING

$$m \leq \hat{m}$$

Figure 25. The Matrix Majorant Is a Bound for the Matrix Block Norm, i.e., the Nonnegative Matrix Each of Whose Elements Is the Norm of the Corresponding Block of a Given Matrix

To apply majorants to dynamical systems, the model is written in the form shown in Figure 26. The matrix A is block diagonal and consists of subsystem dynamics. The subsystem interactions represented by the partitioned matrix G are assumed to be uncertain. By suitable manipulation, uncertainties in the diagonal blocks of A can also be captured by G . By assuming that the spectral norm (largest singular value) of the blocks of G satisfy given bounds, the covariance block-norm inequality is obtained (Figure 27). This inequality is interpreted in the sense of nonnegative matrices, i.e., element-by-element, and $*$ denotes the Hadamard (element-by-element) product.

$$\begin{aligned}\dot{\mathbf{x}} &= (\mathbf{A} + \mathbf{G})\mathbf{x} + \mathbf{w} & \dot{\mathbf{Q}} &= (\mathbf{A} + \mathbf{G})\mathbf{Q} + \mathbf{Q}(\mathbf{A} + \mathbf{G})^T + \mathbf{V}\end{aligned}$$

$$\mathbf{A} = \begin{bmatrix} \mathbf{A}_1 & 0 & \cdots \\ 0 & \mathbf{A}_2 & \cdots \\ \vdots & \vdots & \ddots \end{bmatrix} \quad \mathbf{G} = \begin{bmatrix} 0 & \mathbf{G}_{12} & \cdots \\ \mathbf{G}_{21} & 0 & \cdots \\ \vdots & \vdots & \ddots \end{bmatrix}$$

Known Subsystem Dynamics Uncertain Subsystem Interactions

$$\mathbf{V} = \begin{bmatrix} \mathbf{V}_1 & \mathbf{V}_{12} & \cdots \\ \mathbf{V}_{21} & \mathbf{V}_2 & \cdots \\ \vdots & \vdots & \ddots \end{bmatrix} \quad \mathbf{Q} = \begin{bmatrix} \mathbf{Q}_1 & \mathbf{Q}_{12} & \cdots \\ \mathbf{Q}_{21} & \mathbf{Q}_2 & \cdots \\ \vdots & \vdots & \ddots \end{bmatrix}$$

Noise Intensity State Covariance

Figure 26. The Large-Scale System Model Involves Known Local Dynamics and Uncertain Interactions

$$\begin{aligned}\dot{\mathbf{x}} &= (\mathbf{A} + \mathbf{G})\mathbf{x} + \mathbf{w} & J &= E[\mathbf{x}^T \mathbf{R} \mathbf{x}] = \text{tr } \mathbf{Q} \mathbf{R} \\ \dot{\mathbf{Q}} &= (\mathbf{A} + \mathbf{G})\mathbf{Q} + \mathbf{Q}(\mathbf{A} + \mathbf{G})^T + \mathbf{V} & J &= |\lambda(\mathbf{A}_1 \oplus \mathbf{A}_1)|\end{aligned}$$

$$\mathbf{V} = \begin{bmatrix} \|\mathbf{V}_1\|_F & \|\mathbf{V}_{12}\|_F & \cdots \\ \|\mathbf{V}_{21}\|_F & \|\mathbf{V}_2\|_F & \cdots \\ \vdots & \vdots & \ddots \end{bmatrix} \quad \mathbf{Q} = \begin{bmatrix} \|\mathbf{Q}_1\|_F & \|\mathbf{Q}_{12}\|_F & \cdots \\ \|\mathbf{Q}_{21}\|_F & \|\mathbf{Q}_2\|_F & \cdots \\ \vdots & \vdots & \ddots \end{bmatrix}$$

$$\begin{bmatrix} 0 & \bar{\sigma}(\mathbf{G}_{12}) & \cdots \\ \bar{\sigma}(\mathbf{G}_{21}) & 0 & \cdots \\ \vdots & \vdots & \ddots \end{bmatrix} \leq \xi$$

$$\Downarrow$$

$$\mathbf{A} * \mathbf{Q} \leq \xi \mathbf{Q} + \mathbf{Q} \xi^T + \mathbf{V}$$

Figure 27. The Block-Norm Matrix of the State Covariance Satisfies a Lyapunov-Type Inequality Involving Nonnegative Matrices

To achieve robustness, the covariance block-norm inequality is replaced by the majorant Lyapunov equation (Figure 28). The solution of the majorant Lyapunov equation provides a bound (majorant) for the block norm of the covariance thereby guaranteeing both robust stability and performance.

MAJORANT LYAPUNOV EQUATION

$$A \cdot \hat{Q} = G\hat{Q} + \hat{Q}G^T + V$$

$$\sigma(G_{ij}) \leq G_{ij}$$



$$Q \leq \hat{Q}$$



■ Robust Stability

■ Robust Performance

Figure 28. The Corresponding Nonnegative Matrix Equation Yields a Majorant for the State Covariance and Hence Robust Stability and Performance

It is interesting to note that numerical solution of the majorant Lyapunov equation requires no new techniques. Utilizing properties of M matrices, the solution can be obtained monotonically by means of a straightforward iterative technique (Figure 29).

MLE has a unique solution iff $\{\hat{Q}_K, K=0, 1, \dots, \infty\}$ where:

$$\hat{Q}_0 = 0$$

$$\hat{Q}_{K+1} = H^{-1} \cdot (G \hat{Q}_K + \hat{Q}_K G^T + V)$$

$$(H^{-1})_{mn} \triangleq H^{-1}_{mn}$$

converges. If so, then:

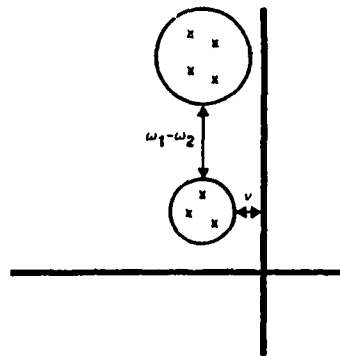
$$\hat{Q} = \lim_{K \rightarrow \infty} \hat{Q}_K$$

$$J - J_0 \leq 2 \sum_{K=1}^T (\text{tr } \hat{P}_K) (G \hat{Q})_{KK}$$

$$(0 = A_K^T \hat{P}_K + \hat{P}_K A_K + R_K)$$

Figure 29. By Exploiting the Properties of M-Matrices, the Majorant Lyapunov Equation Can Be Solved Monotonically by Means of a Simple Iterative Technique

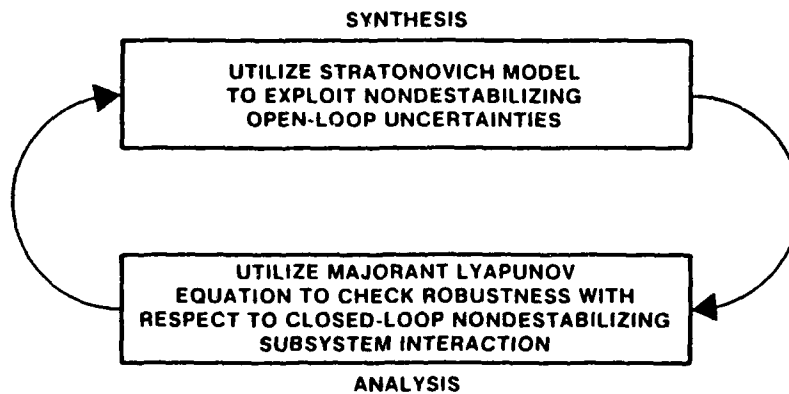
An illustrative application of the majorant Lyapunov equation involves lightly damped subsystems (Figure 30). As shown in [A42] (and expected intuitively), robustness with respect to uncertain subsystem interaction is proportional to the frequency separation between the subsystems. The ability to capture this robustification mechanism is a unique feature of the majorant Lyapunov function not available from quadratic (i.e., scalar) Lyapunov functions or vector Lyapunov functions ([18,19]).



$$\text{Majorant Lyapunov Equation Bound} \sim \nu \sqrt{(2\nu)^2 + (\omega_1 - \omega_2)^2}$$

Figure 30. Robustness Bounds for Uncertain Coupling in Modal Systems Are Proportional to the Frequency Separation Between Subsystems

The next step in the majorant development involves a hierarchy of finer and finer robustness bounds which account for higher order subsystem interactions, e.g., the interaction between the *i*th and *j*th subsystems via the *k*th subsystem. The second member of the hierarchy (Figure 31) provides robustness guarantees with respect to frequency uncertainties. The interesting aspect of this robustness test is the fact that the performance bound is characterized precisely by a Stratonovich model. Hence the Stratonovich model can be viewed as an approximation to a robustness bound, while exploiting the Stratonovich/majorant relationship leads to a natural synthesis/analysis scheme (Figure 32) which nonconservatively exploits open-loop nondestabilizing uncertainties.



Stratonovich synthesis = approximation to majorant analysis

Figure 31. The Stratonovich Synthesis Model Provides a First Approximation to the Majorant Analysis Bounds

Second member of the hierarchy:

$$A \cdot \hat{Q} + \hat{I}[\hat{Q}] = \xi \langle \hat{Q} \rangle + \langle \hat{Q} \rangle \xi^T + V$$

$$J - \text{tr}[\hat{Q}R] \leq 2 \sum_{K=1}^r (\text{tr} \hat{P}_K) (\xi \langle \hat{Q} \rangle)_K K$$

$$0 = A\hat{Q} + \hat{Q}A^T + \hat{I}[\hat{Q}] + V$$

$$0 = A^T \hat{P} + \hat{P}A + \hat{I}[\hat{P}] + R$$

where:

$$\langle \hat{Q} \rangle \triangleq \text{off-diagonal part of } \hat{Q}$$

$$\hat{I}[\cdot] = \text{Stratonovich model operator}$$

- Tighter bound—incorporates more information on A and G
- Predicts stability when $(A + A^T)$ stable, $G = -G^T$
- "Nominal" performance, $\text{tr}[\hat{Q}R]$, given by Stratonovich model

Figure 32. The Refined Majorant Bound Incorporates a Stratonovich Covariance Model

5. Numerical Algorithms and Examples

Practical design of controllers is only possible when efficient, reliable algorithms are available. Indeed, the optimal projection equations are readily solvable and have been applied to a wide variety of examples. Numerical results appear in [A3-A6, A8, A11, A12, A14-A16, A18, A19, A21-A24, A26-A28, A30-A33, A39, A42, A44, A46]. Two distinctly different algorithms have been developed thus far, namely, an iterative method and a homotopy algorithm.

The iterative method, developed in [A14, A16, A44] and further studied in [20, 21], is outlined in Figure 33. The nice feature of this approach is that only a standard LQG software package is required for its implementation. The basic motivation for the method is the observation that the main source of coupling is via the terms involving τ_1 . The coupling is absent, of course, when τ is the identity, i.e., LQG. Note also that the terms involving τ_1 are small when R_2 and V_2 are large, i.e., when control cost is high and the measurement noise is significant. This case, which yields low-authority controllers, is approximately characterized by decoupled control-design and controller-reduction operations. Thus it is not surprising that LQG reduction techniques are most successful when controller authority is low.

Since the τ_1 terms occasion the greatest difficulty, it appears advantageous to bring them into play gradually. This can be accomplished by fixing τ after each iteration to yield updated values of Q , P , \hat{Q} and \hat{P} . Furthermore, τ is introduced gradually by means of α to reduce its rank.

The crucial step of the algorithm concerns the construction of the projection τ from the pseudogramians \hat{Q} and \hat{P} . Specifically, τ can be characterized (see [A22]) as the sum of eigenprojections of $\hat{Q}\hat{P}$, where each choice of eigenprojections may correspond to a local extremal. However, the necessary conditions do not specify which eigenprojections are to be selected for obtaining a particular local solution. Nevertheless, there do exist useful methods for constructing τ . For example, component-cost decomposition methods ([22]) when applied within the optimal projection framework often permit efficient identification of the global optimum.

Although the iterative method is convenient to use because it utilizes readily available software, it is suboptimal in the sense that it does not fully exploit the structure of the equations. Specifically, while the iterative method addresses a system of four $n \times n$ matrix equations, careful analysis reveals that because of the rank deficiency of the projection the problem can be recast as four $n_c \times n_c$ equations. Hence, when n_c is much smaller than n , which is clearly the most

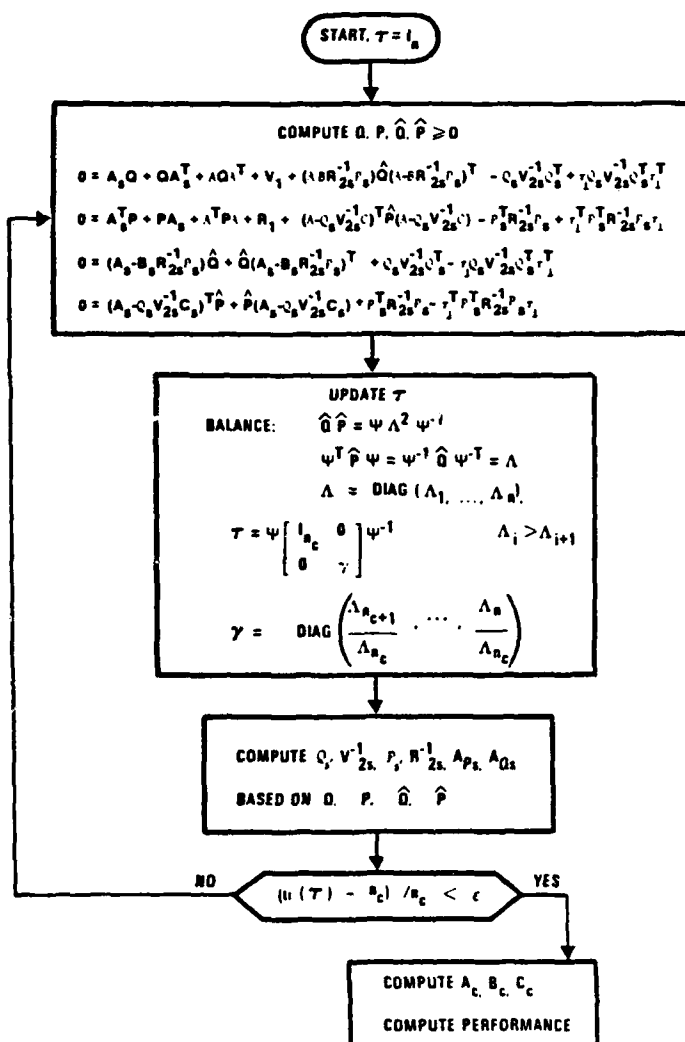


Figure 33. The Iterative Method for Solving the Robustified OPUS Design Equations Requires Only an LQG Software Package and Involves Refinement of the Optimal Projection τ

desirable case for practical implementation, there exists considerable opportunity for increased computational efficiency. Furthermore, and most satisfying, the computational complexity decreases with n_c as is intuitively expected below that required by LQG design. Hence the optimal projection approach has computational complexity less than LQG reduction methods for which LQG is but the first step.

S. Richter ([23,A46]) has developed a homotopy algorithm which fully exploits this crucial structure. Numerical experiments thus far have shown that considerable computational savings can be achieved over the iterative method. Furthermore, by applying topological degree theory to investigate the branches and character of the local extremals, it can be shown that the maximum number of possible extremals is

$$\binom{\min(n,m,l)}{n_c}$$

if $n_c \leq \min(n,m,l)$ or 1 otherwise. Hence in most practical cases the equations support a relatively small number of solutions.

Both the iterative method and the homotopy algorithm have been applied to a design problem involving an 8th-order flexible structure originally due to D. Enns and considered in [1]. Specifically, a variety of LQG reduction methods are compared in [1] for a range of controller authorities. These methods include:

1. Enns: This method is a frequency-weighted, balanced realization technique applicable to either model or controller reduction.
2. Glover: This method utilizes the theory of Hankel norm optimal approximation for controller reduction.
3. Davis and Skelton: This is a modification of compensator reduction via balancing which addresses the case of unstable controllers.
4. Yousuff and Skelton: This is a further modification of balancing for handling stable or unstable controllers.
5. Liu and Anderson: In place of using a balanced approximation of the compensator transfer function directly, this method approximates the component parts of a fractional representation of the compensator.

All of the above methods proceed by first obtaining the full-order LQG compensator design for a high-order state-space model and then reducing the dimension of the resulting LQG compensator.

Figure 34 summarizes the results reported in [1] for the above LQG reduction methods along with results obtained using the iterative method for solving the optimal projection equations. Here q_2 is a scale factor for the plant disturbance noise affecting controller authority. Clearly, LQG reduction methods experience increasing difficulty as authority increases, i.e., as the r_1 terms become increasingly more important in coupling the control and reduction operations. For the low authority cases, the optimal projection calculations, which were performed on a Harris H800 minicomputer, appeared to incur roughly the same computational burden as the LQG reduction methods. Although the optimal projection computational burden increases with authority, comparison with the LQG reduction methods is not meaningful because of the difficulty experienced by these methods in achieving closed-loop stability. See [A44] for further details and for comparisons involving transient response.

The homotopy algorithm was also applied to the example considered in [1]. One of the main goals of the development effort was to extend the range of disturbance intensity or, equivalently, observer bandwidth, out beyond $q_2 = 2000$. To this end, second-order ($n_c = 2$) controllers were obtained with relatively little computation for $q_2 = 10,000$, 100,000 and 1,000,000. In addition, the performance of each reduced-order controller was within 25% of LQG. These cases can surely be expected to present a nontrivial challenge to both the LQG reduction methods and the iterative optimal projection method.

Numerical solution of the robustified optimal projection equations has been carried out for several examples. For illustrative purposes a 2x2 example was considered in [A26] and the results illustrated in Figure 35 indicate performance/robustness tradeoffs possible. The variable left-shift technique was applied in [A19] to the NASA SCOLE problem with frequency uncertainties. The robustness of LQG and two robustified designs is shown in Figure 36. The plots illustrate the degradation in performance due to simultaneous perturbation of all modal frequencies. Note that LQG is rendered unstable by +5% frequency perturbation while a high-authority robustified design improves this region to +8%. The low-authority design increases this region significantly while sacrificing 6% nominal performance.

| Method | $\frac{q_2}{n_c}$ | 0.01 | 0.1 | 1 | 10 | 100 | 1000 | 2000 |
|--------------------|-------------------|------|-----|---|----|-----|------|------|
| Enns | 7 | S | S | S | S | S | S | S |
| | 6 | S | S | S | S | S | S | S |
| | 5 | S | S | S | S | S | S | S |
| | 4 | S | S | S | S | S | S | U |
| | 3 | S | S | S | S | S | S | S |
| | 2 | S | S | S | S | U | U | U |
| Glover | 7 | S | S | S | S | S | U | S |
| | 6 | S | S | S | S | U | U | U |
| | 5 | S | S | S | S | U | U | U |
| | 4 | S | S | S | S | U | U | U |
| | 3 | S | S | U | S | U | U | U |
| | 2 | S | U | S | U | S | U | U |
| Davis & Skelton | 7 | S | U | U | S | S | S | S |
| | 6 | S | S | S | S | S | S | S |
| | 5 | S | U | S | S | S | U | U |
| | 4 | S | S | U | S | S | U | U |
| | 3 | U | U | U | U | U | U | U |
| | 2 | S | U | S | U | U | U | U |
| Yousuff & Skelton | 7 | S | S | S | S | U | U | U |
| | 6 | S | S | S | S | U | U | U |
| | 5 | S | S | S | S | U | U | U |
| | 4 | S | S | S | U | U | U | U |
| | 3 | S | U | U | U | U | U | U |
| | 2 | S | S | S | U | U | U | U |
| Liu & Anderson | 7 | S | S | S | S | S | S | U |
| | 6 | S | S | S | S | S | S | U |
| | 5 | S | S | S | S | S | S | S |
| | 4 | S | S | S | S | S | S | S |
| | 3 | S | S | S | S | S | U | U |
| | 2 | S | S | S | S | S | S | S |
| Optimal Projection | 7 | S | S | S | S | S | S | S |
| | 6 | S | S | S | S | S | S | S |
| | 5 | S | S | S | S | S | S | S |
| | 4 | S | S | S | S | S | S | S |
| | 3 | S | S | S | S | S | S | S |
| | 2 | S | S | S | S | S | S | S |

S - The closed-loop system is stable
 U - The closed-loop system is unstable

Figure 34. The Optimal Projection Approach Was Compared to Several LQG Reduction Techniques Over a Range of Controller Authorities for an Example of Enns

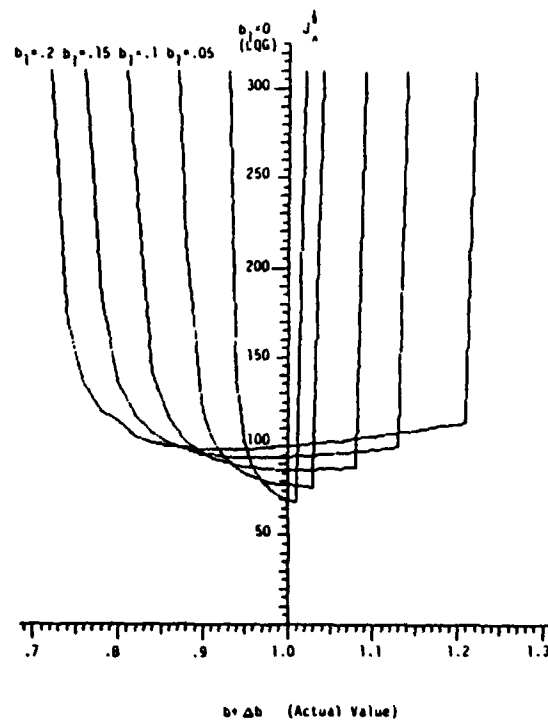


Figure 35. The Robustified Optimal Projection Equations Provide Robustness/Performance Tradeoffs for a Highly Sensitive Nominal LQG Design

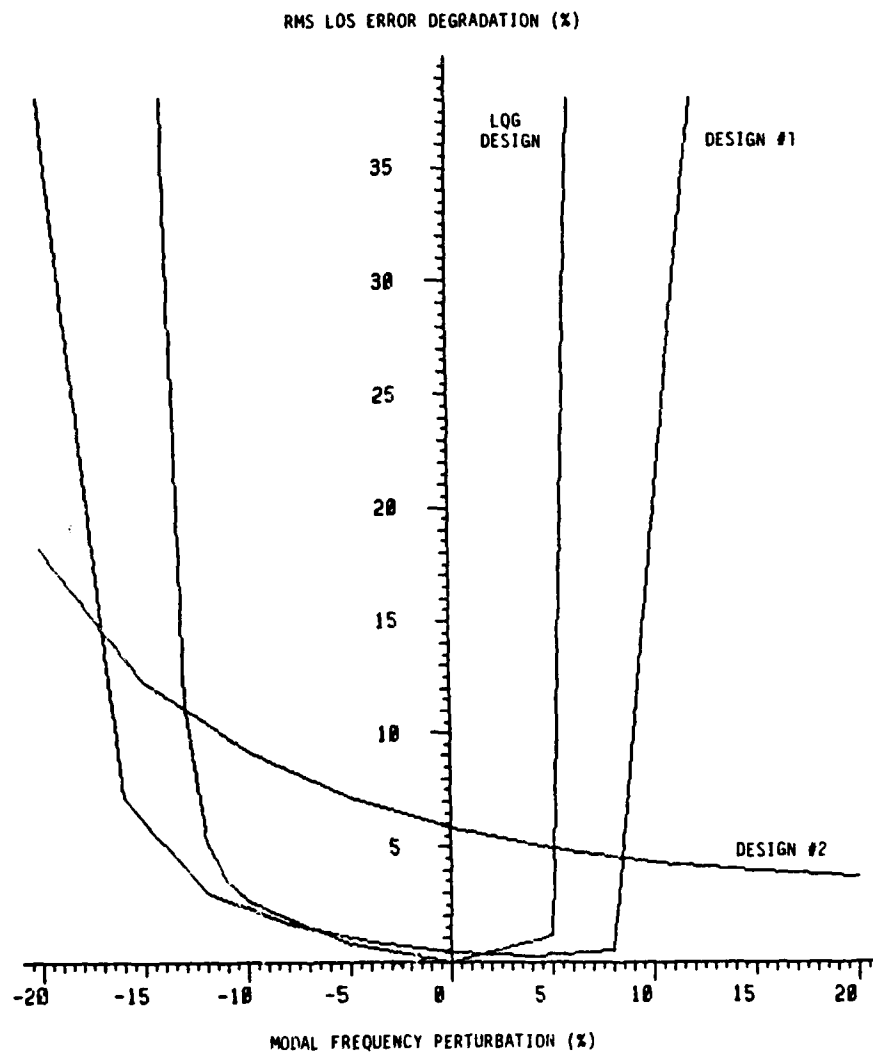


Figure 36. The Stratonovich Model Robustifies the LQG Design for the NASA SCOLE Model with Uncertain Modal Frequencies

6. Additional Extensions

The robustified optimal projection design machinery has been further extended to encompass a larger number of design cases arising in practical application. Here we merely list the extensions:

1. Discrete-time and sampled-data controllers ([A28,A30,A34,A35]).
2. Decentralized controllers ([A39]).
3. Nonstrictly proper controllers ([A37]).
4. Distributed parameter systems ([A25]).

7. Concluding Remarks

The machinery provided by OPUS for designing active controllers for flexible structures has been reviewed. The basic machinery is a system of coupled Riccati and Lyapunov equations which directly generalize the classical LQG result to include both a constraint on controller order and a model of parameter uncertainty. The overall approach thus encompasses all major design tradeoffs arising in vibration-suppression applications. Substantial numerical experience has been gained through an iterative method requiring only an LQG software package and, more recently, by means of a highly efficient homotopy algorithm developed by S. Richter. The overall approach opens the door for effective design of implementable controllers for large precision space structures.

Acknowledgment. We wish to thank Ms. Jill M. Straehle for the excellent preparation of this paper.

General References

1. Y. Liu and B. D. O. Anderson, "Controller Reduction Via Stable Factorization and Balancing," Int. J. Contr., Vol. 44, pp. 507-531, 1986.
2. G. Zames, "Feedback and Optimal Sensitivity: Model Reference Transformations, Multiplicative Seminorms, and Approximate Inverses," IEEE Trans. Autom. Contr., Vol. AC-26, pp. 301-320, 1981.
3. J. C. Doyle, "Guaranteed Margins for LQG Regulators," IEEE Trans. Autom. Contr., Vol. AC-23, pp. 756-757, 1978.
4. J. C. Doyle and G. Stein, "Multivariable Feedback Design: Concepts for a Classical/Modern Synthesis," IEEE Trans. Autom. Contr., Vol. AC-26, pp. 4-16, 1981.
5. G. Stein and M. Athans, "The LQG/LTR Procedure for Multivariable Feedback Control Design," IEEE Trans. Autom. Contr., Vol. AC-32, pp. 105-114, 1987.
6. E. Soroka and U. Shaked, "On the Robustness of LQ Regulators," IEEE Trans. Autom. Contr., Vol. AC-29, pp. 664-665, 1984.
7. U. Shaked and E. Soroka, "On the Stability Robustness of the Continuous-Time LQG Optimal Control," IEEE Trans. Autom. Contr., Vol. AC-30, 1039-1043.
8. J. C. Doyle, "Analysis of Feedback Systems with Structured Uncertainties," IEE Proc., Vol. 129, pp. 242-250, 1982.
9. S. S. L. Chang and T. K. C. Peng, "Adaptive Guaranteed Cost Control of Systems with Uncertain Parameters," IEEE Trans. Autom. Contr., Vol. AC-17, pp. 474-483, 1972.
10. A. Vinkler and L. J. Wood, "Multistep Guaranteed Cost Control of Linear Systems with Uncertain Parameters," J. Guid. Contr., Vol. 2, pp. 449-456, 1979.
11. I. R. Petersen and C. V. Hollot, "A Riccati Equation Approach to the Stabilization of Uncertain Systems," Automatica, Vol. 22, pp. 433-448, 1986.
12. D. Hinrichsen and A. J. Pritchard, "Stability Radius for Structured Perturbations and the Algebraic Riccati Equation," Sys. Contr. Lett., Vol. 8, pp. 105-113, 1986.
13. B. D. O. Anderson and J. B. Moore, Linear Optimal Control, Prentice-Hall, Englewood Cliffs, NJ, 1970.
14. G. N. Milstein, "Design of Stabilizing Controller with Incomplete State Data for Linear Stochastic System with Multiplicative Noise," Autom. Remote Contr., Vol. 43, pp. 653-659, 1982.
15. G. A. Adamian and J. S. Gibson, "Sensitivity of Closed-Loop Eigenvalues and Robustness," preprint.
16. A. M. Ostrowski, "On Some Metrical Properties of Operator Matrices and Matrices Partitioned into Blocks," J. Math. Anal. Appl., Vol. 2, pp. 161-209, 1961.
17. G. Dahlquist, "On Matrix Majorants and Minorants, with Applications to Differential Equations," Lin. Alg. Appl., Vol. 52/53, pp. 199-216, 1983.

18. D. D. Siljak, Large-Scale Dynamic Systems, Elsevier/North-Holland, 1978.
19. M. Ikeda and D. D. Siljak, "Generalized Decompositions of Dynamic Systems and Vector Lyapunov Functions," IEEE Trans. Autom. Contr., Vol. AC-26, pp. 1118-1125, 1981.
20. A. Gruzen, "Robust Reduced Order Control of Flexible Structures," C. S. Draper Laboratory Report #CSDL-T-900, April 1986.
21. A. Gruzen and W. E. Vander Velde, "Robust Reduced-Order Control of Flexible Structures Using the Optimal Projection/Maximum Entropy Design Methodology," AIAA Guid. Nav. Contr. Conf., Williamsburg, VA, August 1986.
22. A. Yousuff and R. E. Skelton, "Controller Reduction by Component Cost Analysis," IEEE Trans. Autom. Contr., Vol. AC-24, pp. 520-530, 1984.
23. S. Richter and R. DeCarlo, "Continuation Methods: Theory and Applications," IEEE Trans. Autom. Contr., Vol. 28, pp. 660-665, 1983.

OPUS References

- A1. D. C. Hyland, "The Modal Coordinate/Radiative Transfer Formulation of Structural Dynamics--Implications for Vibration Suppression in Large Space Platforms," MIT Lincoln Laboratory, TR-27, 14 March 1979.
- A2. D. C. Hyland, "Optimal Regulation of Structural Systems With Uncertain Parameters," MIT Lincoln Laboratory, TR-551, 2 February 1981, DDC# AD-A099111/7.
- A3. D. C. Hyland, "Active Control of Large Flexible Spacecraft: A New Design Approach Based on Minimum Information Modelling of Parameter Uncertainties," Proc. Third VPI&SU/AIAA Symposium, pp. 631-646, Blacksburg, VA, June 1981.
- A4. D. C. Hyland, "Optimal Regulator Design Using Minimum Information Modelling of Parameter Uncertainties: Ramifications of the New Design Approach," Proc. Third VPI&SU/AIAA Symposium, pp. 701-716, Blacksburg, VA, June 1981.
- A5. D. C. Hyland and A. N. Madiwale, "Minimum Information Approach to Regulator Design: Numerical Methods and Illustrative Results," Proc. Third VPI&SU/AIAA Symposium, pp. 101-118, Blacksburg, VA, June 1981.
- A6. D. C. Hyland and A. N. Madiwale, "A Stochastic Design Approach for Full-Order Compensation of Structural Systems with Uncertain Parameters," Proc. AIAA Guid. Contr. Conf., pp. 324-332, Albuquerque, NM, August 1981.
- A7. D. C. Hyland, "Optimality Conditions for Fixed-Order Dynamic Compensation of Flexible Spacecraft with Uncertain Parameters," AIAA 20th Aerospace Sciences Meeting, paper 82-0312, Orlando, FL, January 1982.
- A8. D. C. Hyland, "Structural Modeling and Control Design Under Incomplete Parameter Information: The Maximum Entropy Approach," AFOSR/NASA Workshop in Modeling, Analysis and Optimization Issues for Large Space Structures, Williamsburg, VA, May 1982.
- A9. D. C. Hyland, "Minimum Information Stochastic Modelling of Linear Systems with a Class of Parameter Uncertainties," Proc. Amer. Contr. Conf., pp. 620-627, Arlington, VA, June 1982.

- A10. D. C. Hyland, "Maximum Entropy Stochastic Approach to Control Design for Uncertain Structural Systems," Proc. Amer. Contr. Conf., pp. 680-688, Arlington, VA, June 1982.
- A11. D. C. Hyland, "Minimum Information Modeling of Structural Systems with Uncertain Parameters," Proceedings of the Workshop on Applications of Distributed System Theory to the Control of Large Space Structures, G. Rodriguez, ed., pp. 71-88, JPL, Pasadena, CA, July 1982.
- A12. D. C. Hyland and A. N. Madiwale, "Fixed-Order Dynamic Compensation Through Optimal Projection," Proceedings of the Workshop on Applications of Distributed System Theory to the Control of Large Space Structures, G. Rodriguez, ed., pp. 409-427, JPL, Pasadena, CA, July 1982.
- A13. D. C. Hyland, "Mean-Square Optimal Fixed-Order Compensation--Beyond Spillover Suppression," paper 1403, AIAA Astrodynamics Conference, San Diego, CA, August 1982.
- A14. D. C. Hyland, "The Optimal Projection Approach to Fixed-Order Compensation: Numerical Methods and Illustrative Results," AIAA 21st Aerospace Sciences Meeting, paper 83-0303, Reno, NV, January 1983.
- A15. D. C. Hyland, "Mean-Square Optimal, Full-Order Compensation of Structural Systems with Uncertain Parameters," MIT Lincoln Laboratory, TR-626, 1 June 1983.
- A16. D. C. Hyland, "Comparison of Various Controller-Reduction Methods: Suboptimal Versus Optimal Projection," Proc. AIAA Dynamics Specialists Conf., pp. 381-389, Palm Springs, CA, May 1984.
- A17. D. C. Hyland and D. S. Bernstein, "The Optimal Projection Equations for Fixed-Order Dynamic Compensation," IEEE Trans. Autom. Contr., Vol. AC-29, pp. 1034-1037, 1984.
- A18. D. C. Hyland, "Application of the Maximum Entropy/Optimal Projection Control Design Approach for Large Space Structures," Proc. Large Space Antenna Systems Technology Conference, pp. 617-654, NASA Langley, December 1984.
- A19. L. D. Davis, D. C. Hyland and D. S. Bernstein, "Application of the Maximum Entropy Design Approach to the Spacecraft Control Laboratory Experiment (SCOLE)," Final Report, NASA Langley, January 1985.
- A20. D. S. Bernstein and D. C. Hyland, "The Optimal Projection Equations for Reduced-Order State Estimation," IEEE Trans. Autom. Contr., Vol. AC-30, pp. 583-585, 1985.
- A21. D. S. Bernstein and D. C. Hyland, "Optimal Projection/Maximum Entropy Stochastic Modelling and Reduced-Order Design Synthesis," Proc. IFAC Workshop on Model Error Concepts and Compensation, Boston, MA, June 1985, pp. 47-54, R. E. Skelton and D. H. Owens, eds., Pergamon Press, Oxford, 1986.
- A22. D. C. Hyland and D. S. Bernstein, "The Optimal Projection Equations for Model Reduction and the Relationships Among the Methods of Wilson, Skelton and Moore," IEEE Trans. Autom. Contr., Vol. AC-30, pp. 1201-1211, 1985.
- A23. D. S. Bernstein and D. C. Hyland, "The Optimal Projection/Maximum Entropy Approach to Designing Low-Order, Robust Controllers for Flexible Structures," Proc. 24th IEEE Conf. Dec. Contr., pp. 745-752, Fort Lauderdale, FL, December 1985.

- A24. D. S. Bernstein, L. D. Davis, S. W. Greeley and D. C. Hyland, "Numerical Solution of the Optimal Projection/Maximum Entropy Design Equations for Low-Order, Robust Controller Design," Proc. 24th IEEE Conf. Dec. Contr., pp. 1795-1798, Fort Lauderdale, FL, December 1985.
- A25. D. S. Bernstein and D. C. Hyland, "The Optimal Projection Equations for Finite-Dimensional Fixed-Order Dynamic Compensation of Infinite-Dimensional Systems," SIAH J. Contr. Optim., Vol. 24, pp. 122-151, 1986.
- A26. D. S. Bernstein and S. W. Greeley, "Robust Controller Synthesis Using the Maximum Entropy Design Equations," IEEE Trans. Autom. Contr., Vol. AC-31, pp. 362-364, 1986.
- A27. D. C. Hyland, D. S. Bernstein, L. D. Davis, S. W. Greeley and S. Richter, "MEOP: Maximum Entropy/Optimal Projection Stochastic Modelling and Reduced-Order Design Synthesis," Final Report, Air Force Office of Scientific Research, Bolling AFB, Washington, DC, April 1986.
- A28. D. S. Bernstein, L. D. Davis and D. C. Hyland, "The Optimal Projection Equations for Reduced-Order, Discrete-Time Modelling, Estimation and Control," J. Guid. Contr. Dyn., Vol. 9, pp. 288-293, 1986.
- A29. D. S. Bernstein and S. W. Greeley, "Robust Output-Feedback Stabilization: Deterministic and Stochastic Perspectives," Proc. Amer. Contr. Conf., pp. 1818-1826, Seattle, WA, June 1986.
- A30. D. S. Bernstein, L. D. Davis and S. W. Greeley, "The Optimal Projection Equations for Fixed-Order, Sampled-Data Dynamic Compensation with Computation Delay," IEEE Trans. Autom. Contr., Vol. AC-31, pp. 859-862, 1986.
- A31. D. S. Bernstein, "OPUS: Optimal Projection for Uncertain Systems," Annual Report, Air Force Office of Scientific Research, Bolling AFB, Washington, DC, October 1986.
- A32. B. J. Boan and D. C. Hyland, "The Role of Metal Matrix Composites for Vibration Suppression in Large Space Structures," Proc. NAC Spacecraft Survivability Tech. Conf., NMCIAC Kaman Tempo Publ., Stanford Research Institute, Palo Alto, CA, October 1986.
- A33. D. C. Hyland, "An Experimental Testbed for Validation of Control Methodologies in Large Space Optical Structures," SPIE Optoelectronics and Laser Applications Conference, Los Angeles, CA, January 1987.
- A34. W. M. Haddad and D. S. Bernstein, "The Optimal Projection Equations for Discrete-Time Reduced-Order State Estimation for Linear Systems with Multiplicative White Noise," Sys. Contr. Lett., 1987.
- A35. D. S. Bernstein and W. M. Haddad, "The Optimal Projection Equations for Discrete-Time Fixed-Order Dynamic Compensation of Linear Systems with Multiplicative White Noise," Int. J. Contr., 1987.
- A36. W. M. Haddad and D. S. Bernstein, "The Optimal Projection Equations for Reduced-Order State Estimation: The Singular Measurement Noise Case," IEEE Trans. Autom. Contr., 1987.
- A37. D. S. Bernstein, "The Optimal Projection Equations for Static and Dynamic Output Feedback: The Singular Case," IEEE Trans. Autom. Contr., 1987.

- A38. D. S. Bernstein and D. C. Hyland, "The Optimal Projection Equations for Reduced-Order Modelling, Estimation and Control of Linear Systems with Multiplicative White Noise," J. Optim. Thy. Appl., 1987.
- A39. D. S. Bernatein, "Sequential Design of Decentralized Dynamic Compensators Using the Optimal Projection Equations," Int. J. Contr., 1987.
- A40. D. S. Bernstein and W. M. Haddad, "Optimal Output Feedback for Nonzero Set Point Regulation," Proc. Amer. Contr. Conf., Minneapolis, MN, June 1987.
- A41. D. S. Bernatein, "Robust Static and Dynamic Output-Feedback Stabilization: Deterministic and Stochastic Perspectives," IEEE Trans. Autom. Contr., 1987.
- A42. D. C. Hyland and D. S. Bernstein, "The Majorant Lyapunov Equation: A Nonnegative Matrix Equation for Guaranteed Robust Stability and Performance of Large Scale Systems," IEEE Trans. Autom. Contr., 1987.
- A43. D. S. Bernatein and W. M. Haddad, "The Optimal Projection Equations with Petersen-Hollot Bounds: Robust Controller Synthesis with Guaranteed Structured Stability Radius," submitted.
- A44. S. W. Greeley and D. C. Hyland, "Reduced-Order Compensation: LQG Reduction Versus Optimal Projection," submitted.
- A45. W. M. Haddad, Robust Optimal Projection Control-System Synthesis, Ph.D. Dissertation, Department of Mechanical Engineering, Florida Institute of Technology, Melbourne, FL, March 1987.
- A46. S. Richter, "A Homotopy Algorithm for Solving the Optimal Projection Equations for Fixed-Order Dynamic Compensation: Existence, Convergence and Global Optimality," Proc. Amer. Contr. Conf., Minneapolis, MN, June 1987.

APPENDIX B: Fixed-Structure Design

32. D. C. Hyland and D. S. Bernstein, "The Optimal Projection Equations for Model Reduction and the Relationships Among the Methods of Wilson, Skelton and Moore," IEEE Trans. Autom. Contr., Vol. AC-30, pp. 1201-1211, 1985.
29. D. S. Bernstein and D. C. Hyland, "The Optimal Projection Equations for Reduced-Order State Estimation," IEEE Trans. Autom. Contr., Vol. AC-30, pp. 583-585, 1985.
24. D. C. Hyland and D. S. Bernstein, "The Optimal Projection Equations for Fixed-Order Dynamic Compensation," IEEE Trans. Autom. Contr., Vol. AC-29, pp. 1034-1037, 1984.

The Optimal Projection Equations for Model Reduction and the Relationships Among the Methods of Wilson, Skelton, and Moore

DAVID C. HYLAND AND DENNIS S. BERNSTEIN, MEMBER, IEEE

Abstract—First-order necessary conditions for quadratically optimal reduced-order modeling of linear time-invariant systems are derived in the form of a pair of modified Lyapunov equations coupled by an oblique projection which determines the optimal reduced-order model. This form of the necessary conditions considerably simplifies previous results of Wilson [1] and clearly demonstrates the quadratic extremality and nonoptimality of the balancing method of Moore [2]. The possible existence of multiple solutions of the optimal projection equations is demonstrated and a relaxation-type algorithm is proposed for computing these local extrema. A component-cost analysis of the model-error criterion similar to the approach of Skelton [3] is utilized at each iteration to direct the algorithm to the global minimum.

I. INTRODUCTION

THE problem of approximating a high-order linear dynamical system with a relatively simpler system, i.e., the model-reduction problem, has received considerable attention in recent years. Among the myriad papers devoted to this problem are the notable contributions of Wilson [1], Moore [2], and Skelton [3] with which the present paper is concerned. In his 1970 paper, Wilson proposed an optimality-based approach to model reduction which involves minimizing the steady-state, quadratically weighted¹ output error when the original system and reduced-order model are subjected to white-noise inputs. For the resulting parameter-optimization problem, he obtained first-order

necessary conditions which have the form of an aggregation (as, e.g., [4]) and which involve the solution of two Lyapunov equations each of order $n + n_m$, where n and n_m are the orders of the original and reduced-order models, respectively [5], [6].

Some time later, Moore proposed a quite different approach to model reduction based upon system-theoretic arguments as opposed to optimality criteria. Using the eigenvalues of the product of the controllability and observability gramians (which satisfy $n \times n$ Lyapunov equations), his method identifies subsystems which contribute little to the impulse response of the overall system. Such "weak" subsystems are thus eliminated to obtain a reduced-order model. This technique, known as balancing, has been vigorously developed in the recent literature [7]–[11]. Since this approach is completely independent of optimality considerations, there is, of course, no expectation that such reduced-order models are in any sense optimal.

A third approach to model reduction, proposed by Skelton [3], [12], also utilizes a quadratic optimality criterion as in [1]. However, rather than proceeding from necessary conditions as does Wilson, Skelton determines for a given basis the contribution (cost) of each state in a decomposition of the error criterion and truncates those with the least value. Although this approach is guided by optimality considerations, no rigorous guarantee of optimality is possible because of dependence on the choice of state space basis.

The present paper has five main objectives, the first of which is to show how the complex optimality conditions of Wilson can be transformed without loss of generality into much simpler and more tractable forms. The transformation is facilitated by exploiting the presence of an oblique (i.e., nonorthogonal) projection which was not recognized in [1]² and which arises as a direct consequence of optimality. The resulting "optimal projection equations" constitute a coupled system of two $n \times n$

Manuscript received October 2, 1984; revised February 1, 1985. This paper is based on a prior submission of March 14, 1984. Paper recommended by Past Associate Editor, B. R. Barmish. This work was supported in part by the Air Force at Lincoln Laboratory/M.I.T., Lexington, MA.

The authors are with Harris Corporation, Government Aerospace Systems Division, Controls Analysis and Synthesis Group, Melbourne, FL 32901.

¹ The quadratic error criterion has been chosen for consideration in the present paper because of its relation to the standard engineering practice of stating specifications in terms of rms deviation.

² The projection was, however, pointed out in [28, p. 29].

modified Lyapunov equations [see (2.13), (2.14) or (2.21), (2.22)] whose solutions are given by a pair of rank- n_m controllability and observability pseudogramians. The highly structured form of these equations gives crucial insight into the set of local extrema satisfying the first-order necessary conditions.

The second objective of the paper is to show how the optimal projection equations provide a rigorous extremality context for Moore's balancing method and to clearly demonstrate its quadratic nonoptimality. Although for some problems the "weak subsystem" hypothesis leads to a nearly optimal reduced-order model, we construct examples for which the reduced-order model obtained from the balancing method is much worse with respect to the least-squares criterion than the quadratically optimal reduced-order model. In general, all that can be said is that the presence of a weak subsystem indicates that the reduced-order model obtained by truncation in the balanced basis may be in the proximity of an extremal of the quadratically optimal model-reduction problem; however, this extremal may very well be a global maximum. It should be noted that in a recent paper [13] Kabamba has used bounds on the model error to demonstrate the quadratic nonoptimality of the balancing method.

The third objective of the paper is to demonstrate via an example the mechanism responsible for the existence of multiple extrema of the optimal model-reduction problem. By characterizing the optimal projection as a sum of rank-1 eigenprojections of the product of the rank-deficient pseudogramians, it is immediately clear that the first-order necessary conditions of the problem are ambiguous in the sense that they fail to specify which n_m eigenprojections comprise the optimal projection corresponding to a solution (i.e., global minimum) of the optimal model-reduction problem. Specifically, since the pseudogramians can be rank deficient in $\binom{n}{n_m} = n!/n_m!(n - n_m)!$ ways, there may be precisely this many extremal projections corresponding to an identical number of local extrema.

The fourth objective of the paper is to propose a numerical algorithm for solving the optimal projection equations by exploiting their structure and taking advantage of the available insights. By expressing the modified Lyapunov equations in the form of "standard" Lyapunov equations, an iterative relaxation-type algorithm is developed. The crucial aspect of the proposed algorithm involves extracting an oblique projection at each step from the product of the solutions of the Lyapunov equations. Since $\binom{n}{n_m}$ rank- n_m projections can be extracted from the product of two $n \times n$ positive-definite matrices, it is quickly evident that the criterion by which the n_m eigenprojections are chosen determines which of the numerous local extrema will be reached. If, for example, the projection is chosen in accordance with the n_m largest eigenvalues of the product of the solutions of the Lyapunov equations, then it should not be surprising in view of the previous discussion that a global maximum may very well be reached. In this case, the first iteration of this algorithm involves Lyapunov equations whose solutions are the controllability and observability gramians and the eigenvalues in question are precisely the squares of the second-order modes [2, p. 24]. Thus, the first iteration coincides with the (nonoptimal) balancing approach of [2].

Since the optimal projection equations are a consequence of differential (local) properties, it should not be expected that they alone would possess the inherent ability to identify the global minimum. Moreover, because of the number of local extrema, second-order necessary conditions appear to be useless. Instead, we investigate an approach which chooses the eigenprojections according to a component-cost analysis of the model-error criterion. This technique can lead to a global minimum by effectively eliminating the local extrema which have considerably greater cost than the global minimum. This approach is philosophically identical to the component cost analysis of Skelton [3], [12]. Essentially, then, component cost analysis is utilized at each iteration to direct the algorithm to the global minimum. Although our application of this technique is admittedly heuristic, it should be noted that it is essentially proposed as a device for efficiently

"sorting out" the local extrema which satisfy the otherwise mathematically rigorous necessary conditions. Hence, we propose component cost analysis as a crucial step in bridging the gap between local extremality and global optimality.

It should be pointed out that neither the numerical algorithm proposed in this paper nor the iterative algorithm developed in [4] and [5] has been proven to be convergent. The *principal contribution* of the present paper, however, is not a particular proposed algorithm but rather the *revelations concerning the structure of the first-order necessary conditions*. The proposed numerical algorithm should be considered but a prelude to a full investigation into numerical algorithms for the optimal projection equations. It should also be noted that the presence of the optimal projection was not exploited in developing the iterative algorithms in [4] and [5] (in fact, it did not even appear in [1]) and hence crucial insight into local extrema was lacking.

The fifth and last objective of the paper is to point out the connection between the optimal projection equations for model reduction obtained herein and the first-order necessary conditions obtained recently for two closely related problems, namely, reduced-order state estimation and fixed-order dynamic compensation.

The plan of the paper is as follows. Section II begins with general notation and definitions followed by the model-reduction problem statement and the main theorem which presents the optimal projection equations for model reduction. A series of remarks considers various aspects of the main theorem and sets the stage for discussing connections with [1] and [2]. Section III contains a detailed discussion of the sense in which the optimal projection equations simplify the necessary conditions given in [1], and Section IV shows how the approach of [2] is approximately extremal. Section V presents a simple example which clearly displays the possible existence of multiple extrema satisfying the optimal projection equations. This example shows that the balancing method of [2] may lead to a nonoptimal reduced-order model and suggests a heuristic procedure for selecting the eigenprojections comprising the projection corresponding to the global minimum, i.e., the *optimal projection*. In Section VI, a numerical algorithm for solving the optimal projection equations is proposed and applied to an illustrative example considered previously in [1] and [2] as well as to some interesting examples considered recently by Kabamba in [13]. Related results on reduced-order dynamic compensation and state estimation are briefly reviewed in Section VII and suggestions for further research are given in Section VIII. The proof of the main theorem appears in the Appendix.

II. PROBLEM STATEMENT AND MAIN RESULT

The following notation and definitions will be used throughout the paper:

| | |
|--|--|
| I_r | $r \times r$ identity matrix |
| Z^T | transpose of vector or matrix Z |
| Z^{-T} | $(Z^T)^{-1}$ or $(Z^{-1})^T$ |
| $\rho(Z)$ | rank of matrix Z |
| $\text{tr } Z$ | trace of square matrix Z |
| $\ Z\ $ | $[\text{tr } Z Z^T]^{1/2}$ |
| Z_{ij} | (i, j) -element of matrix Z |
| $\text{diag } (\alpha_1, \dots, \alpha_r)$ | $r \times r$ diagonal matrix with listed diagonal elements |
| E_i | matrix with unity in the (i, i) position and zeros elsewhere |
| ξ | expected value |
| R, R^{*s} | real numbers, $r \times s$ real matrices |
| stable matrix | matrix with eigenvalues in open left half plane |
| nonnegative-definite matrix | symmetric matrix with nonnegative eigenvalues |
| positive-definite matrix | symmetric matrix with positive eigenvalues |

| | |
|-------------------------------|---|
| semisimple matrix | matrix similar to a diagonal matrix [14, p. 10] |
| nonnegative semisimple matrix | matrix similar to a nonnegative-definite matrix |
| positive-semisimple matrix | matrix similar to a positive-definite matrix |
| positive-diagonal matrix | diagonal matrix with positive diagonal elements |
| n, m, l, n_m | positive integers, $1 \leq n_m \leq n$ |
| x, u, y, x_m, y_m | n, m, l, n_m, l -dimensional vectors |
| A, B, C | $n \times n, n \times m, l \times n$ matrices |
| A_m, B_m, C_m | $n_m \times n_m, n_m \times m, l \times n_m$ matrices |
| R, V | $l \times l, m \times m$ positive-definite matrices |

We consider the following problem.

Optimal Model-Reduction Problem: Given the controllable and observable system

$$\dot{x} = Ax + Bu, \quad (2.1)$$

$$y = Cx \quad (2.2)$$

find a reduced-order model

$$\dot{x}_m = A_m x_m + B_m u, \quad (2.3)$$

$$y_m = C_m x_m \quad (2.4)$$

which minimizes the quadratic model-reduction criterion³

$$J(A_m, B_m, C_m) \triangleq \lim_{t \rightarrow \infty} \mathbb{E}[(y - y_m)^T R (y - y_m)],$$

where the input $u(t)$ is white noise with positive-definite intensity V . To guarantee that J is finite, it is assumed that A is stable and we restrict our attention to the set of admissible reduced-order models

$$\mathcal{Q} \triangleq \{(A_m, B_m, C_m) : A_m \text{ is stable}\}.$$

Since the value of J is independent of the internal realization of the transfer function corresponding to (2.3) and (2.4), we further restrict our attention to the set

$$\mathcal{Q}_+ \triangleq \{(A_m, B_m, C_m) \in \mathcal{Q} :$$

$$(A_m, B_m) \text{ is controllable and } (A_m, C_m) \text{ is observable}\}.$$

The following lemma is needed for the statement of the main result.

Lemma 2.1: Suppose $\hat{Q}, \hat{P} \in \mathbb{R}^{n \times n}$ are nonnegative definite. Then $\hat{Q}\hat{P}$ is nonnegative semisimple. Furthermore, if $\rho(\hat{Q}\hat{P}) = n_m$ then there exist $G, \Gamma \in \mathbb{R}^{n_m \times n}$ and positive-semisimple $M \in \mathbb{R}^{n_m \times n_m}$ such that

$$\hat{Q}\hat{P} = G^T M \Gamma, \quad (2.5)$$

$$\Gamma G^T = I_{n_m}. \quad (2.6)$$

Proof: By [14, Theorem 6.2.5, p. 123], there exists $n \times n$ invertible Φ such that the nonnegative-definite matrices $D_Q \triangleq \Phi \hat{Q} \Phi^T$ and $D_P \triangleq \Phi^{-T} \hat{P} \Phi^{-1}$ are both diagonal. Hence, $D_Q D_P$ is nonnegative definite and $\hat{Q}\hat{P} = \Phi^{-1} D_Q D_P \Phi$ is nonnegative semisimple. Next introduce $n \times n$ orthogonal U to effect a rearrangement of basis if necessary so that

$$\hat{Q}\hat{P} = \Phi \begin{bmatrix} \Lambda & 0 \\ 0 & 0 \end{bmatrix} \Phi^{-1},$$

³ J will occasionally be referred to as the "model-reduction error" or, simply, as the "cost."

where $\Phi \triangleq \Phi U$ and $n_m \times n_m$ Λ is positive diagonal. Hence, for all $n_m \times n_m$ invertible S ,

$$\hat{Q}\hat{P} = \Phi \begin{bmatrix} S & \\ 0 & \end{bmatrix} (S^{-1} \Lambda S) [S^{-1} \ 0] \Phi^{-1}$$

and thus, (2.5) and (2.6) hold with $G = [S^T \ 0] \Phi^T$, $M = S^{-1} \Lambda S$ and $\Gamma = [S^{-1} \ 0] \Phi^{-1}$. ■

For convenience in stating the main theorem, we shall refer to $G, \Gamma \in \mathbb{R}^{n_m \times n}$ and positive-semisimple $M \in \mathbb{R}^{n_m \times n_m}$ satisfying (2.5) and (2.6) as a (G, M, Γ) -factorization of $\hat{Q}\hat{P}$. Also, define the positive-definite controllability and observability gramians

$$W_c \triangleq \int_0^\infty e^{At} B V B^T e^{A^T t} dt,$$

$$W_o \triangleq \int_0^\infty e^{A^T t} C^T R C e^{A t} dt,$$

which satisfy the dual Lyapunov equations

$$0 = A W_c + W_c A^T + B V B^T, \quad (2.7)$$

$$0 = A^T W_o + W_o A + C^T R C. \quad (2.8)$$

Main Theorem: Suppose $(A_m, B_m, C_m) \in \mathcal{Q}_+$ solves the optimal model-reduction problem. Then there exist nonnegative-definite matrices $\hat{Q}, \hat{P} \in \mathbb{R}^{n \times n}$ such that, for some (G, M, Γ) -factorization of $\hat{Q}\hat{P}$, A_m, B_m , and C_m are given by

$$A_m = \Gamma A G^T, \quad (2.9)$$

$$B_m = \Gamma B, \quad (2.10)$$

$$C_m = C G^T, \quad (2.11)$$

and such that, with $\tau \triangleq G^T \Gamma$, the following conditions are satisfied:

$$\rho(\hat{Q}) = \rho(\hat{P}) = \rho(\hat{Q}\hat{P}) = n_m, \quad (2.12)$$

$$0 = \tau[A\hat{Q} + \hat{Q}A^T + BVB^T], \quad (2.13)$$

$$0 = [A^T \hat{P} + \hat{P}A + C^T R C] \tau. \quad (2.14)$$

Several comments are in order. First, note that the main theorem consists of necessary conditions in the form of two modified Lyapunov equations (2.13) and (2.14) plus rank conditions (2.12) which must possess nonnegative-definite solutions \hat{Q}, \hat{P} when an optimal reduced-order model exists. We shall call \hat{Q} and \hat{P} the *controllability* and *observability pseudogramians*, respectively, since they are analogous to W_c and W_o and yet have rank deficiency. The modified Lyapunov equations are coupled by the $n \times n$ matrix τ which is a projection (idempotent matrix) since

$$\tau^2 = G^T \Gamma G^T \Gamma = G^T I_{n_m} \Gamma = \tau.$$

Note that, in general, τ is an *oblique* projection and not necessarily an orthogonal projection since it may not be symmetric. We shall refer to a projection τ corresponding to a solution (i.e., global minimum) of the optimal model-reduction problem as an "optimal projection." It should be stressed that the form of the optimal reduced-order model (2.7)-(2.9) is a direct consequence of optimality and not the result of an *a priori* assumption on the structure of the reduced-order model.

Since the optimal projection equations are first-order necessary conditions for optimality, they may possess multiple solutions corresponding to various local extrema such as local maxima, local minima, saddle points, etc. The following definition will prove useful.

Definition 2.1: Nonnegative-definite $\hat{Q}, \hat{P} \in \mathbb{R}^{n \times n}$ are *extremal* if (2.12)-(2.14) are satisfied. $(A_m, B_m, C_m) \in \mathcal{Q}_+$ is

extremal if there exist extremal \hat{Q}, \hat{P} such that (A_m, B_m, C_m) is given by (2.9)–(2.11) for some (G, M, Γ) -factorization of $\hat{Q}\hat{P}$. The corresponding projection τ is an extremal projection.

Proposition 2.1: Suppose (A_m, B_m, C_m) is extremal. Then the model-reduction error is given by⁴

$$J(A_m, B_m, C_m) = 2\text{tr}[(\hat{Q}\hat{P} - W_c W_o)A]. \quad (2.15)$$

Proof: The proof is given at the end of Appendix A. ■

Remark 2.1: Noting the identities

$$-2\text{tr}[W_c W_o A] = \text{tr}[C^T R C W_c] = \text{tr}[B V B^T W_o], \quad (2.16)$$

which follow from (2.7) and (2.8), (2.15) can be written for extremal (A_m, B_m, C_m) as

$$\begin{aligned} J(A_m, B_m, C_m) &= 2\text{tr}[\hat{Q}\hat{P}A] + \text{tr}[C^T R C W_c] \\ &= 2\text{tr}[\hat{Q}\hat{P}A] + \text{tr}[B V B^T W_o]. \end{aligned} \quad (2.17)$$

For convenience in the following discussion, let $\hat{Q}, \hat{P}, G, M, \Gamma$, and τ correspond to some extremal (A_m, B_m, C_m) . Now observe that if x_m is replaced by Sx_m , where S is an arbitrary nonsingular matrix, then an "equivalent" reduced-order model is obtained with (A_m, B_m, C_m) replaced by $(SA_m S^{-1}, SB_m, C_m S^{-1})$. Since $J(A_m, B_m, C_m) = J(SA_m S^{-1}, SB_m, C_m S^{-1})$, one would expect the main theorem to apply also to $(SA_m S^{-1}, SB_m, C_m S^{-1})$. Indeed, the following result shows that this transformation corresponds to the alternative factorization $\hat{Q}\hat{P} = (S^{-T}G)^T(SMS^{-1})(S\Gamma)$ and, moreover, that all (G, M, Γ) -factorizations of $\hat{Q}\hat{P}$ are related by an invertible transformation.

Proposition 2.2: If $S \in \mathbb{R}^{n_m \times n_m}$ is invertible, then $\hat{G} = S^{-T}G$, $\hat{\Gamma} = S\Gamma$ and $\hat{M} = SMS^{-1}$ satisfy

$$\hat{Q}\hat{P} = \hat{G}^T \hat{M} \hat{\Gamma}, \quad (2.5)'$$

$$\hat{\Gamma} \hat{G}^T = I_{n_m}. \quad (2.6)'$$

Conversely, if $\hat{G}, \hat{\Gamma} \in \mathbb{R}^{n_m \times n}$ and invertible $\hat{M} \in \mathbb{R}^{n_m \times n_m}$ satisfy (2.5)' and (2.6)', then there exists invertible $S \in \mathbb{R}^{n_m \times n_m}$ such that $\hat{G} = S^{-T}G$, $\hat{\Gamma} = S\Gamma$ and $\hat{M} = SMS^{-1}$.

Proof: The first part is immediate. The second part follows by taking $S \triangleq \hat{M}^{-1} \hat{\Gamma} \hat{G}^T M$, noting $S^{-1} = M \hat{G}^T \hat{M}^{-1}$ and using the identities $\hat{\Gamma} \hat{G}^T M \hat{G}^T = \hat{M}$ and $M \hat{G}^T = \hat{\Gamma} \hat{G}^T \hat{M}$. ■

The next result shows that there exists a similarity transformation which simultaneously diagonalizes $\hat{Q}\hat{P}$ and τ .

Proposition 2.3: There exists invertible $\Phi \in \mathbb{R}^{n \times n}$ such that

$$\hat{Q} = \Phi^{-1} \begin{bmatrix} \Lambda_Q & 0 \\ 0 & 0 \end{bmatrix} \Phi^{-T}, \quad \hat{P} = \Phi^T \begin{bmatrix} \Lambda_P & 0 \\ 0 & 0 \end{bmatrix} \Phi, \quad (2.18)$$

$$\hat{Q}\hat{P} = \Phi^{-1} \begin{bmatrix} \Lambda & 0 \\ 0 & 0 \end{bmatrix} \Phi, \quad \tau = \Phi^{-1} \begin{bmatrix} I_{n_m} & 0 \\ 0 & 0 \end{bmatrix} \Phi, \quad (2.19a, b)$$

where $\Lambda_Q, \Lambda_P \in \mathbb{R}^{n_m \times n_m}$ are positive diagonal, $\Lambda \triangleq \Lambda_Q \Lambda_P$ and the diagonal elements of Λ are the eigenvalues of M . Consequently,

$$\hat{Q} = \tau \hat{Q}, \quad \hat{P} = \hat{P} \tau. \quad (2.20)$$

Proof: By [14, Theorem 6.2.5, p. 123], and by (2.12), there exists $n \times n$ invertible Φ such that (2.18) holds and thus (2.19a) also holds. Define

$$\hat{G} = [I_{n_m} \ 0] \Phi^{-T}, \quad \hat{M} = \Lambda \quad \text{and} \quad \hat{\Gamma} = [I_{n_m} \ 0] \Phi$$

so that (2.5)' and (2.6)' are satisfied. By the second part of Proposition 2.2 there exists invertible $S \in \mathbb{R}^{n_m \times n_m}$ such that $G =$

$S^T \hat{G}$, $M = S^{-1} \hat{M} S$ and $\Gamma = S^{-1} \hat{\Gamma}$. Now (2.19b) follows from

$$\tau = G^T \Gamma = \hat{G}^T \hat{\Gamma} = \Phi^{-1} \begin{bmatrix} I_{n_m} & 0 \\ 0 & 0 \end{bmatrix} \Phi. \quad \blacksquare$$

It is useful to present an alternative form of the optimal model-reduction equations (2.13) and (2.14). For convenience, define the notation

$$\tau_\perp \triangleq I_n - \tau.$$

Proposition 2.4: Equations (2.13) and (2.14) are equivalent, respectively, to

$$0 = A \hat{Q} + \hat{Q} A^T + B V B^T - \tau_\perp B V B^T \tau_\perp^T, \quad (2.21)$$

$$0 = A^T \hat{P} + \hat{P} A + C^T R C - \tau_\perp^T C^T R C \tau_\perp. \quad (2.22)$$

Proof: By (2.20), (2.21) = (2.13) + (2.13) τ + (2.13) τ and (2.13) = τ (2.21). Similarly, (2.14) and (2.22) are equivalent. ■

Remark 2.2: Noting the identities

$$-2 \text{tr}[\hat{Q}\hat{P}A] = \text{tr}[C^T R C \hat{Q}] = \text{tr}[B V B^T \hat{P}], \quad (2.23)$$

which follow from (2.20)–(2.22), (2.17) can be written for extremal (A_m, B_m, C_m) as

$$J(A_m, B_m, C_m) = \text{tr}[C^T R C (W_c - \hat{Q})] = \text{tr}[B V B^T (W_o - \hat{P})]. \quad (2.24)$$

To facilitate the discussion in the following sections, we consider the change of basis $\hat{x} \triangleq \Phi x$, where Φ is given by Proposition 2.3. Writing (2.1) and (2.2) as

$$\dot{\hat{x}} = \hat{A} \hat{x} + \hat{B} u, \quad (2.25)$$

$$y = \hat{C} \hat{x}, \quad (2.26)$$

where

$$\hat{A} \triangleq \Phi A \Phi^{-1}, \quad \hat{B} \triangleq \Phi B, \quad \hat{C} \triangleq C \Phi^{-1},$$

(2.9)–(2.11) become

$$A_m = \hat{\Gamma} \hat{A} \hat{G}^T, \quad (2.27)$$

$$B_m = \hat{\Gamma} \hat{B}, \quad (2.28)$$

$$C_m = \hat{C} \hat{G}^T, \quad (2.29)$$

where

$$\hat{\Gamma} \triangleq \Gamma \Phi^{-1}, \quad \hat{G} \triangleq G \Phi^T$$

satisfy

$$\hat{G}^T \hat{\Gamma} = \begin{bmatrix} I_{n_m} & 0 \\ 0 & 0 \end{bmatrix}, \quad \hat{\Gamma} \hat{G}^T = I_{n_m}. \quad (2.30)$$

Note that (2.30) implies

$$\hat{\Gamma} = [S \ 0], \quad \hat{G} = [S^{-T} \ 0], \quad (2.31)$$

for some $n_m \times n_m$ invertible S . Partitioning

$$\hat{x} = \begin{bmatrix} \hat{x}_m \\ \hat{x}_2 \end{bmatrix}, \quad \hat{A} = \begin{bmatrix} \hat{A}_{m1} & \hat{A}_{m2} \\ \hat{A}_{2m} & \hat{A}_{22} \end{bmatrix},$$

$$\hat{B} = \begin{bmatrix} \hat{B}_m \\ \hat{B}_2 \end{bmatrix}, \quad \hat{C} = [\hat{C}_m \ \hat{C}_2],$$

where $\hat{x}_m \in \mathbb{R}^{n_m}$ and \hat{A}_m, \hat{B}_m and \hat{C}_m are $n_m \times n_m$, $n_m \times m$ and

⁴ The expressions (2.15)–(2.17) and (2.23)–(2.24) will be used in Sections V and VI.

$l \times n_m$, respectively, (2.27)–(2.29) and (2.31) yield

$$A_m = S\hat{A}_m S^{-1}, \quad B_m = S\hat{B}_m, \quad C_m = \hat{C}_m S^{-1}.$$

This shows that the optimal reduced-order model (modulo a state transformation) can be obtained by truncating the last $n - n_m$ states of the original system when it is expressed in the basis with respect to which \hat{Q} and \hat{P} have the diagonal forms

$$\begin{bmatrix} \Lambda_Q & 0 \\ 0 & 0 \end{bmatrix} \text{ and } \begin{bmatrix} \Lambda_P & 0 \\ 0 & 0 \end{bmatrix}.$$

Since the optimal projection τ has the simple form

$$\begin{bmatrix} I_{n_m} & 0 \\ 0 & 0 \end{bmatrix}$$

in this basis, we shall refer to (2.25) and (2.26) as an *optimal projection realization* of (2.1) and (2.2). Note that when (2.21) and (2.22) are expanded in an optimal projection basis (i.e., a basis corresponding to an optimal projection realization) they assume the form

$$0 = \hat{A}_m \Lambda_Q + \Lambda_Q \hat{A}_m^T + \hat{B}_m V \hat{B}_m^T, \quad (2.32)$$

$$0 = \hat{A}_{2m} \Lambda_Q + \hat{B}_2 V \hat{B}_m^T, \quad (2.33)$$

$$0 = \hat{A}_m^T \Lambda_P + \Lambda_P \hat{A}_m + \hat{C}_m^T R \hat{C}_m, \quad (2.34)$$

$$0 = \Lambda_P \hat{A}_{m2} + \hat{C}_m^T R \hat{C}_2. \quad (2.35)$$

If Φ in Proposition 2.23 is replaced by

$$\begin{bmatrix} (\Lambda_Q^{-1} \Lambda_P)^{1/4} & 0 \\ 0 & I_{n-n_m} \end{bmatrix} \Phi,$$

which corresponds to a change of basis for the reduced-order model obtained by truncation, then Λ_Q and Λ_P are both replaced by $(\Lambda_Q \Lambda_P)^{1/2}$ and hence this can be called a *balanced optimal projection basis*, utilizing the terminology of [2]. Thus, in a balanced optimal projection realization, Λ_Q and Λ_P appearing in (2.32)–(2.35) are equal.

The next result provides an interesting closed-form characterization of an extremal projection in terms of the Drazin generalized inverse of $\hat{Q}\hat{P}$. Since $(\hat{Q}\hat{P})^2 = G^T M^2 \Gamma$, and hence $\rho(\hat{Q}\hat{P})^2 = \rho(\hat{Q}\hat{P})$, the "index" of $\hat{Q}\hat{P}$ (see [15, p. 121]) is 1. In this case, the Drazin inverse is traditionally called the group inverse and is denoted by $(\hat{Q}\hat{P})^\#$ [15, p. 124]. Since, as is easily verified, $(\hat{Q}\hat{P})^\# = G^T M^{-1} \Gamma$, (2.6) leads to the following result.

Proposition 2.5: An extremal projection τ is given by

$$\tau = \hat{Q}\hat{P}(\hat{Q}\hat{P})^\#. \quad (2.36)$$

An alternative representation for an extremal projection will prove useful for developing a numerical algorithm for solving (2.21) and (2.22). If $Q, P \in \mathbb{R}^{n \times n}$ are nonnegative definite then by Lemma 2.1 QP is nonnegative semisimple and thus there exists invertible $\Psi \in \mathbb{R}^{n \times n}$ such that

$$QP = \Psi^{-1} \Omega \Psi,$$

where $\Omega = \text{diag}(\omega_1, \dots, \omega_r)$ and $\omega_i \geq 0$ are the eigenvalues of QP . Now define the i th eigenprojection [16, p. 41]

$$\Pi_i[QP] \triangleq \Psi^{-1} E_i \Psi,$$

which is a rank-1 oblique projection. Note that QP has the decomposition

$$QP = \sum_{i=1}^r \omega_i \Pi_i[QP].$$

Proposition 2.6: An extremal projection τ is given by

$$\tau = \sum_{i=1}^{n_m} \Pi_i[\hat{Q}\hat{P}]. \quad (2.37)$$

where the i th eigenprojection $\Pi_i[\hat{Q}\hat{P}]$ corresponds to the i th nonzero eigenvalue λ_i of $\hat{Q}\hat{P}$.

III. RELATIONSHIP TO WILSON'S FORM OF THE NECESSARY CONDITIONS

The optimal model-reduction problem considered in the previous section is identical to the problem considered by Wilson in [1] with the minor exception that he sets $R = I_r$. In [1] G and Γ are denoted by θ_1^T and θ_1 , (2.6) appears as (15), and (2.9)–(2.11) are given by (14a, b). Note that in [1], θ_1 and θ_2 depend upon the solutions of a pair of $(n + n_m) \times (n + n_m)$ Lyapunov equations [see (7), (9) of [1] or (A.2), (A.3) of the present paper] whose coefficients and nonhomogeneous terms depend in turn on A_m , B_m , and C_m [see (A.10)–(A.15)]. The advantage of the $n \times n$ optimal projection equations (2.21) and (2.22) over the form of the necessary conditions given in [1] [see (A.10)–(A.15)] is that the optimal projection equations are *independent* of A_m , B_m , and C_m . Hence, this permits the development of numerical algorithms which avoid the need to choose starting values for A_m , B_m , and C_m . To see this, note that although the unknowns A_m , B_m , and C_m appear explicitly in (A.10)–(A.15), all data in the optimal projection equations (2.13) and (2.14) are known except for the solutions \hat{Q} and \hat{P} . Moreover, the optimal projection τ , which was not recognized in [1], can be seen to play a fundamental role by coupling the modified Lyapunov equations (2.21) and (2.22) and determining (since $\tau = G^T \Gamma$) A_m , B_m , and C_m in (2.7)–(2.9).

IV. RELATIONSHIP TO MOORE'S BALANCING METHOD

In contrast to Wilson's method for model reduction which is based on optimality principles, the approach due to Moore [2] relies on system-theoretic ideas. The main thrust of this approach "is to eliminate any weak subsystem which contributes little to the impulse response matrix" [2, p. 26]. The concept of a "weak subsystem" is defined by means of a dominance relation [2, p. 28] involving similarity invariants called second-order modes. Moore evaluates reduced-order models obtained in this way by computing the relative error in the impulse response given for MIMO systems by [2, p. 29]

$$\epsilon(A_m, B_m, C_m) \triangleq \left[\int_0^\infty \|H_e(t)\|^2 dt / \int_0^\infty \|H(t)\|^2 dt \right]^{1/2},$$

where $H_e(t) \triangleq H(t) - H_m(t)$, $H(t) \triangleq R^{1/2} C e^{A t} B V^{1/2}$ and $H_m(t) \triangleq R^{1/2} C_m e^{A_m t} B_m V^{1/2}$. To discuss this approach in the context of the optimal model-reduction problem, we assume that $V = I_m$ and $R = I_r$.

Proposition 4.1: Suppose $(A_m, B_m, C_m) \in \mathcal{Q}$. Then

$$\begin{aligned} \epsilon(A_m, B_m, C_m) &= \left[-\frac{1}{2} J(A_m, B_m, C_m) / \text{tr}(W_c W_o A) \right]^{1/2} \\ &= [J(A_m, B_m, C_m) / \text{tr}(C^T R C W_c)]^{1/2} \\ &= [J(A_m, B_m, C_m) / \text{tr}(B V B^T W_o)]^{1/2}. \end{aligned} \quad (4.1)$$

Proof: The result follows from (A.1), (A.8), and (A.9) which hold without regard to either optimality or extremality. ■

Note that Proposition 4.1 shows that the relative error in the impulse response is minimized precisely when $J(A_m, B_m, C_m)$ is minimized. Actually, this result is to be expected since, as shown in [1], J can be obtained alternatively by taking $u(t)$ to be an impulse at $t = 0$.

To draw interesting comparisons with the results of [2], choose

$n \times n$ invertible Ψ such that $\Psi W_c \Psi^T$ and $\Psi^{-T} W_o \Psi^{-1}$ are both diagonal and hence

$$W_c W_o = \Psi^{-1} \Sigma^2 \Psi, \quad (4.2)$$

where $\Sigma \triangleq \text{diag}(\sigma_1, \dots, \sigma_n)$ and the second-order modes σ_i (i.e., the positive square roots of the eigenvalues of $W_c W_o$) satisfy $\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_n > 0$. This transformation corresponds to replacing (2.1), (2.2) by

$$\dot{\bar{x}} = \bar{A}\bar{x} + \bar{B}u, \quad (4.3)$$

$$y = \bar{C}\bar{x}, \quad (4.4)$$

where

$$\bar{x} \triangleq \Psi x, \quad \bar{A} \triangleq \Psi A \Psi^{-1}, \quad \bar{B} \triangleq \Psi B, \quad \bar{C} \triangleq C \Psi^{-1}. \quad (4.5)$$

The transformed system (4.3), (4.4), called a principal axis realization [17], can further be chosen so that

$$\Psi W_c \Psi^T = \Psi^{-T} W_o \Psi^{-1} = \Sigma, \quad (4.6)$$

i.e., the balanced realization. Using (4.5), (2.7) and (2.8) become

$$0 = \bar{A}\Sigma + \Sigma\bar{A}^T + \bar{B}\bar{V}\bar{B}^T, \quad (4.7)$$

$$0 = \bar{A}^T\Sigma + \Sigma\bar{A} + \bar{C}^T\bar{R}\bar{C}. \quad (4.8)$$

The model-reduction procedure suggested in [2] involves partitioning

$$\bar{x} = \begin{bmatrix} \bar{x}_m \\ \bar{x}_2 \end{bmatrix}, \quad \bar{A} = \begin{bmatrix} \bar{A}_{m1} & \bar{A}_{m2} \\ \bar{A}_{2m} & \bar{A}_{22} \end{bmatrix},$$

$$\bar{B} = \begin{bmatrix} \bar{B}_m \\ \bar{B}_2 \end{bmatrix}, \quad \bar{C} = [\bar{C}_m \quad \bar{C}_2],$$

where $\bar{x}_m \in \mathbb{R}^{n_m}$ and \bar{A}_m , \bar{B}_m , and \bar{C}_m have corresponding dimension, and extracting the reduced-order model $(\bar{A}_m, \bar{B}_m, \bar{C}_m)$. Hence, the reduced-order model $(\bar{A}_m, \bar{B}_m, \bar{C}_m)$ is extracted from (4.3), (4.4) in essentially the same way the optimal reduced-order model (A_m, B_m, C_m) is extracted from (2.25), (2.26). To see how the optimal-projection realization compares to a principal-axis realization, first note that (2.13) and (2.14) are satisfied by $\bar{Q} = W_c$ and $\bar{P} = W_o$ when the rank conditions (2.10) are ignored. Indeed, since W_c and W_o are positive definite, the rank conditions (2.12) do not hold. If, however, the system (2.1), (2.2) is expressed in the balanced coordinate system (4.3), (4.4) (so that $W_c = W_o = \Sigma$), then the assumption $\sigma_{n_m} \gg \sigma_{n_m+1}$ implies that $\rho(W_c)$, $\rho(W_o)$ and $\rho(W_c W_o)$ are "approximately" equal to n_m and thus, in this sense, condition (2.10) is satisfied. This observation leads to the suggestion that when $\sigma_{n_m} \gg \sigma_{n_m+1}$, W_c and W_o are approximations to solutions \bar{Q} and \bar{P} of the optimal projection equations and the reduced-order model $(\bar{A}_m, \bar{B}_m, \bar{C}_m)$ of Moore is an approximation to some extremal (A_m, B_m, C_m) . There is no guarantee, of course, that any particular extremum corresponds to the global minimum, or even to a local minimum.

V. EXISTENCE OF MULTIPLE EXTREMA AND COMPONENT-COST RANKING

In this section, we show by means of a simple example that the optimal projection equations may possess nonunique solutions corresponding to multiple extrema, e.g., local minima or maxima. We also show how decomposing the cost can identify the global minimum from among the numerous extrema. To begin, let $m = l = n$, $R = V = I_n$,

$$A \triangleq \text{diag}(-\alpha_1, \dots, -\alpha_n),$$

where $\alpha_i > 0$, $i = 1, \dots, n$, and suppose B and C are such that

$$BB^T = \text{diag}(\beta_1, \dots, \beta_n), \quad C^T C = \text{diag}(\gamma_1, \dots, \gamma_n),$$

where $\beta_i > 0$, $\gamma_i > 0$, $i = 1, \dots, n$. Hypothesizing diagonal solutions \bar{Q} and \bar{P} of (2.21) and (2.22) leads to

$$\bar{Q}_{ii} = \frac{\beta_i}{2\alpha_i} \delta_i, \quad \bar{P}_{ii} = \frac{\gamma_i}{2\alpha_i} \delta_i,$$

where each δ_i , $i = 1, \dots, n$ is either zero or one and exactly n_m of the δ_i 's are equal to one. Hence $\tau = \text{diag}(\delta_1, \dots, \delta_n)$. Note that there are $\binom{n}{n_m}$ such solutions of the optimal projection equations corresponding to $\binom{n}{n_m}$ local extrema.

Since

$$W_c = -\frac{1}{2} A^{-1} B B^T, \quad W_o = -\frac{1}{2} A^{-1} C^T C, \quad \bar{Q} = \tau W_c, \quad \bar{P} = \tau W_o$$

and A , W_c , and W_o commute, (2.15) becomes

$$J(A_m, B_m, C_m) = -\frac{1}{2} \text{tr } \tau A^{-1} B B^T C^T C.$$

Hence,

$$J(A_m, B_m, C_m) = \sum_{i=1}^n \xi_i (1 - \delta_i), \quad (5.1)$$

where

$$\xi_i \triangleq \beta_i \gamma_i / 2\alpha_i.$$

To minimize J , it is clear that δ_i should be chosen to be unity for the largest n_m elements of the set $\{\xi_i\}_{i=1}^n$ and zero otherwise. Although this choice is not necessarily unique, it does yield a global minimum. Note that choosing $\delta_i = 1$ is equivalent to selecting a particular eigenprojection $\Pi_i[W_c W_o]$ corresponding to the eigenvalue $\beta_i \gamma_i / 4\alpha_i^2$.

Remark 5.1: The expression in (5.1) can be regarded as a decomposition of the cost in terms of the state variables. The idea of deleting states based on their "component costs" is precisely the "component cost analysis" approach of Skelton [3], [12].

Using the example, it is easy to see that the balancing method of [2], which selects eigenprojections based upon the magnitude of the eigenvalues of $W_c W_o$, i.e., the (squares of the) second-order modes, may yield a grossly suboptimal reduced-order model. To this end, let

$$\alpha_1 = 1, \quad \alpha_2 = 10^6, \quad \beta_1 = 1, \quad \beta_2 = 10^6, \quad \gamma_1 = 1, \quad \gamma_2 = 10^3$$

so that

$$\xi_1 = 0.5, \quad \xi_2 = 500.$$

Clearly, J is minimized ($J = \xi_1$) by choosing $\delta_1 = 0$, $\delta_2 = 1$, which corresponds to truncating the first state variable. If, however, the method of [2] is utilized, then judging by the second-order modes

$$\sigma_1 = 0.5, \quad \sigma_2 = (2.5)^{1/2} \cdot 10^{-2} \approx 0.012,$$

the second state variable should be deleted. This, however, corresponds to choosing $\delta_1 = 1$, $\delta_2 = 0$ with the higher cost $J = \xi_2$. The fact that the balancing approach of [2] fails to determine a solution of the optimal model-reduction problem should not be surprising in view of the fact that the error criterion plays no role in the balancing technique.

Although the above solution exploited the simple structure of this example, it is clear that choosing the global minimum from among the local extrema involves an eigenprojection decomposi-

tion of the cost J . To extend this idea to more general systems, we invoke the following heuristic approximation.

Approximation 5.1: Let Ψ define the balanced basis as in (4.6). Then Ψ also approximately defines a balanced optimal projection basis, i.e.,

$$\Psi \hat{Q} \Psi^T = \Psi^{-T} \hat{P} \Psi^{-1} = \bar{\tau} \Sigma^2, \quad (5.2)$$

where extremal

$$\bar{\tau} \triangleq \Psi \tau \Psi^{-1} = \text{diag}(\delta_1, \dots, \delta_n) \quad (5.3)$$

and

$$\delta_i \in \{0, 1\}, \quad \sum_{i=1}^n \delta_i = n_m.$$

Proposition 5.1: If Approximation 5.1 holds for extremal (A_m, B_m, C_m) then, with $\bar{\tau}_\perp \triangleq I_n - \bar{\tau}$,

$$\begin{aligned} J(A_m, B_m, C_m) &\approx -2\text{tr}[\bar{\tau}_\perp \Sigma^2 \bar{A}] \\ &= 2 \sum_{i=1}^n -\sigma_i^2 \bar{A}_{ii} (1 - \delta_i). \end{aligned} \quad (5.4)$$

Remark 5.2: From (4.7) and (4.8), it follows that (5.4) can be written either as

$$\begin{aligned} J(A_m, B_m, C_m) &= \text{tr}[\bar{\tau}_\perp \Sigma \bar{B} \bar{V} \bar{B}^T] \\ &= \sum_{i=1}^n \sigma_i (B \bar{V} \bar{B}^T)_{ii} (1 - \delta_i) \end{aligned} \quad (5.5)$$

or

$$\begin{aligned} J(A_m, B_m, C_m) &= \text{tr}[\bar{\tau}_\perp \Sigma \bar{C}^T \bar{R} \bar{C}] \\ &= \sum_{i=1}^n \sigma_i (\bar{C}^T \bar{R} \bar{C})_{ii} (1 - \delta_i). \end{aligned} \quad (5.6)$$

Hence, Approximation 5.1 leads to the following component-cost ranking (again, in the sense of Skelton [3], [12]) of the (n_m) extrema satisfying the optimal projection equations.

Component-Cost Ranking: Assume Approximation 5.1 is valid and choose the eigenprojections comprising extremal $\bar{\tau}$ such that

$$\begin{aligned} \delta_i &= 1, & \text{if } -\sigma_i \text{ is among the } n_m \\ & & \text{largest elements of the set } \{-\sigma_i^2 \bar{A}_{ii}\}_{i=1}^n; \\ \delta_i &= 0, & \text{otherwise.} \end{aligned}$$

For comparison purposes, we shall also consider the following ranking of the eigenprojections based upon the eigenvalues of $W_c W_o$ (i.e., second-order modes).

Eigenvalue Ranking: Choose the eigenprojections comprising extremal $\bar{\tau}$ such that

$$\begin{aligned} \delta_i &= 1, & \text{if } -\sigma_i^2 \bar{A}_{ii} \text{ is among the } n_m \\ & & \text{largest elements of the set } \{-\sigma_i\}_{i=1}^n; \\ \delta_i &= 0, & \text{otherwise.} \end{aligned}$$

Remark 5.3: The observation that the second-order modes alone may be a poor guide to determining an optimal reduced-order model has recently been made in [13] where bounds on the model-error criterion were given involving both the second-order modes and suitable weights called *balanced gains*. It can be seen

in Proposition 5.1 that the role of balanced gains in our approach is played by the elements $-\sigma_i \bar{A}_{ii}$ when Approximation 5.1 holds. It can also be seen that the balanced gains of Kabamba yield bounds on the component costs of Skelton.

VI. NUMERICAL SOLUTION OF THE OPTIMAL PROJECTION EQUATIONS

Insofar as the ultimate aim of any model-reduction technique is to permit the development of numerical procedures for reducing high-order models, the optimal projection equations, comprising a coupled system of modified Lyapunov equations, appear promising in this regard. Therefore, we present an iterative computational algorithm that exploits the structure of these equations and the available insights. The reader is strongly reminded that the proposed algorithm is but a first attempt at solving these new equations and alternative algorithms may yet be devised. The basis of this algorithm is the ability to write the modified Lyapunov equations (2.21), (2.22) in the form of "standard" Lyapunov equations (6.1), (6.2) such that the pseudogramians \hat{Q} and \hat{P} are extracted at the final step (6.6). It follows from (2.32)–(2.35) that (2.21), (2.22) are indeed equivalent to (6.1), (6.2) (with $k = \infty$) and (6.6).

Algorithm:

Step 1) Initialize $\tau^{(0)} = I_n$.

Step 2) Solve for $\hat{Q}^{(k)}, \hat{P}^{(k)}$

$$0 = (A - \tau^{(k)} A \tau^{(k)}) \hat{Q}^{(k)} + \hat{Q}^{(k)} (A - \tau^{(k)} A \tau^{(k)})^T + B \bar{V} \bar{B}^T, \quad (6.1)$$

$$0 = (A - \tau^{(k)} A \tau^{(k)})^T \hat{P}^{(k)} + \hat{P}^{(k)} (A - \tau^{(k)} A \tau^{(k)}) + C^T \bar{R} C. \quad (6.2)$$

Step 3) Balance

$$\Phi^{(k)} \hat{Q}^{(k)} (\Phi^{(k)})^T = (\Phi^{(k)})^T \hat{P}^{(k)} (\Phi^{(k)})^{-1} = \Sigma^{(k)}, \quad (6.3)$$

$$\Sigma^{(k)} = \text{diag}(\sigma_1^{(k)}, \dots, \sigma_n^{(k)}), \quad \sigma_1^{(k)} \geq \sigma_2^{(k)} \geq \dots \geq \sigma_n^{(k)} \geq 0.$$

Step 4) If $k > 1$ check for convergence

$$e_k \triangleq \left[\frac{\text{tr}(C^T \bar{R} C W_c) - \text{tr}(C^T \bar{R} C \tau^{(k)} \hat{Q}^{(k)} (\tau^{(k)})^T)}{\text{tr}(C^T \bar{R} C W_c)} \right]^{1/2}. \quad (6.4)$$

If $|e_k - e_{k-1}| < \text{tolerance}$ then go to step 8); else continue;

Step 5) Select n_m eigenprojections

$$\Pi_{i_1}[\hat{Q}^{(k)} \hat{P}^{(k)}], \dots, \Pi_{i_{n_m}}[\hat{Q}^{(k)} \hat{P}^{(k)}],$$

$$\Pi_i[\hat{Q}^{(k)} \hat{P}^{(k)}] \triangleq \Phi^{(k)} E_i (\Phi^{(k)})^{-1}.$$

Step 6) Update

$$\tau^{(k+1)} = \sum_{i=1}^{n_m} \Pi_i[\hat{Q}^{(k)} \hat{P}^{(k)}]. \quad (6.5)$$

Step 7) Check for convergence; if not, increment k and return to Step 2).

Step 8) Set

$$\hat{Q} = \tau^{(\infty)} \hat{Q}^{(\infty)} \tau^{(\infty)T}, \quad \hat{P} = (\tau^{(\infty)})^T \hat{P}^{(\infty)}. \quad (6.6)$$

For convenience, we shall adopt the notation $(A^{(k)}, B^{(k)}, C^{(k)})$, where $k > 0$, to denote the reduced-order model obtained as a result of applying the projection $\tau^{(k)}$, and we define (see Section IV)

$$e_k \triangleq e(A_m^{(k)}, B_m^{(k)}, C_m^{(k)}),$$

i.e., the relative error associated with $(A_m^{(k)}, B_m^{(k)}, C_m^{(k)})$. Note that, in general, $e_k \neq e_k$ since e_k denotes the relative error only for an extremum, i.e., when convergence has been reached.

It should be clear from the discussion in the previous section that the crucial step of the algorithm is Step 5)—the choice of the eigenprojections. For the examples which follow, we shall invoke consistently at Step 5) either the component-cost ranking based upon Approximation 5.1 or the eigenvalue ranking.

Remark 6.1: Note that in the special case $R = I_m$ and $V = I_l$, the first iteration of the algorithm yields $\hat{Q}^{(0)} = W_c$, $\hat{P}^{(0)} = W_o$. If, at Step 5), we choose $i_r = r$, $r = 1, \dots, n_m$, i.e., the eigenprojections are selected according to the eigenvalue ranking, then $(A^{(1)}, B^{(1)}, C^{(1)})$ is precisely the reduced-order model obtained from balancing.

We shall first consider the following example which was treated by both Wilson and Moore. In this example, and those that follow, assume $R = I_m$, $V = I_l$.

Example 6.1:

$$A = \begin{bmatrix} 0 & 0 & 0 & -150 \\ 1 & 0 & 0 & -245 \\ 0 & 1 & 0 & -113 \\ 0 & 0 & 1 & -19 \end{bmatrix}, \quad B = \begin{bmatrix} 4 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \quad C = [0 \ 0 \ 0 \ 1].$$

Table I summarizes the results obtained for the three cases $n_m = 3, 2, 1$ utilizing the eigenvalue ranking. In each case, the proposed algorithm converged linearly in less than eight iterations and, in each case, improvement is evident over previously published results. As pointed out in [2], Wilson's result seems to imply a lack of final convergence. For this example, the balancing approach yields a reduced-order model close to the global minimum.

We now turn to a pair of interesting examples considered in [13].

Example 6.2:

$$A = \begin{bmatrix} -0.005 & -0.99 \\ -0.99 & -5000 \end{bmatrix}, \quad B = \begin{bmatrix} 1 \\ 100 \end{bmatrix}, \quad C = B^T.$$

Table II summarizes the results obtained using the eigenvalue ranking and Table III gives the results when the component-cost ranking is used. It is clear that the former method directs the algorithm to the global maximum whereas the latter approach yields the global minimum.

Example 6.3:

$$A = \begin{bmatrix} -0.25 & -0.4 \\ -0.4 & -0.72 \end{bmatrix}, \quad B = \begin{bmatrix} 1 \\ 1.2 \end{bmatrix}, \quad C = B^T.$$

Table IV reports the results obtained using either the component-cost ranking or the eigenvalue ranking which agree for this example. If the *alternative* eigenprojection is selected then, as expected, the algorithm converges to a global maximum (see Table V). The interesting aspect of this example, as discussed in [13], is that the error $\epsilon_1 = 0.5245$ (see [13]) for the reduced-order model obtained by either eigenprojection ranking is actually *greater* than $\epsilon_1 = 0.3849$ obtained by choosing the *alternative* reduced-order model. This situation seems to indicate that proper eigenprojection selection based upon a cost decomposition is able to direct the algorithm to the global minimum in cases for which the starting values are not nearby.

VII. THE OPTIMAL PROJECTION EQUATIONS FOR FIXED-ORDER DYNAMIC COMPENSATION AND REDUCED-ORDER STATE ESTIMATION

We briefly discuss the relationship between the optimal projection equations for model reduction and analogous results for reduced-order control and estimation problems.

Fixed-Order Dynamic-Compensation Problem: Given the controlled system

$$\dot{x} = Ax + Bu + w_1, \quad (7.1)$$

$$y = Cx + w_2, \quad (7.2)$$

TABLE I
RELATIVE ERROR $\epsilon_m = \epsilon_n$

| Order n_m | Wilson [1] | Moore [2] | Optimal Projection Equations |
|-------------|------------|-----------|------------------------------|
| 3 | — | 0.001311 | 0.001306 |
| 2 | 0.04097 | 0.03938 | 0.03929 |
| 1 | — | 0.4321 | 0.4268 |

TABLE II
EXAMPLE 6.2 WITH EIGENVALUE RANKING

| k | ϵ_k |
|-----|--------------|
| 1 | 0.9950371897 |
| 2 | 0.9950371691 |
| 3 | 0.9950371690 |

TABLE III
EXAMPLE 6.2 WITH COMPONENT-COST RANKING

| k | ϵ_k |
|-----|--------------|
| 1 | 0.0995037 |
| 2 | 0.0995449 |
| 3 | 0.0995924 |
| 4 | 0.0996520 |
| 5 | 0.0997346 |
| 6 | 0.0998648 |
| 7 | 0.1001125 |
| 8 | 0.1007724 |
| 9 | 0.1054569 |
| 10 | 0.0982006 |
| 11 | 0.0975409 |
| 12 | 0.0975342 |
| 13 | 0.0975330 |
| 14 | 0.0975329 |

TABLE IV
EXAMPLE 6.3 USING EITHER RANKING

| k | ϵ_k |
|-----|--------------|
| 1 | 0.646996 |
| 2 | 0.418341 |
| 3 | 0.220994 |
| 4 | 0.177276 |
| 5 | 0.176576 |

TABLE V
EXAMPLE 6.3 WITH THE OPPOSITE RANKING

| k | ϵ_k |
|-----|--------------|
| 1 | 0.7624928516 |
| 2 | 0.9999999961 |
| 3 | 0.9999999975 |
| . | . |
| . | . |
| 29 | 0.9999999999 |

design a fixed-order dynamic compensator

$$\dot{x}_c = A_c x_c + B_c y, \quad (7.3)$$

$$u = C_c x_c \quad (7.4)$$

which minimizes the performance criterion

$$J(A_c, B_c, C_c) \triangleq \lim_{t \rightarrow \infty} \mathbb{E}[x^T R_1 x + u^T R_2 u], \quad (7.5)$$

where $u \in \mathbb{R}^m$, $x_c \in \mathbb{R}^{n_c}$, $n_c \leq n$, w_1 is white disturbance noise, w_2 is nonsingular white observation noise, R_1 is nonnegative definite, and R_2 is positive definite.

Necessary conditions characterizing optimal (A_c, B_c, C_c) have been developed in [18]–[22] along the same lines as the main theorem. These conditions, called the optimal projection equations for fixed-order dynamic compensation, consist of four matrix equations (two modified Riccati equations and two modified Lyapunov equations) coupled by a projection. The modified Riccati equations, not surprisingly, are similar in form to the covariance and cost Riccati equations of LQG theory and the modified Lyapunov equations are similar to the optimal model-reduction equations (2.13) and (2.14). Hence, while the modified Riccati equations govern optimal estimation and optimal control, the additional modified Lyapunov equations characterize "optimal reduction." The important fact that all four equations are coupled supports the view that optimal fixed-order dynamic compensators *cannot*, in general, be designed by means of a stepwise procedure, e.g., by either open-loop model reduction followed by LQG or LQG followed by closed-loop model reduction.

Midway between the model-reduction and fixed-order dynamic-compensation problems lies the following problem.

Reduced-Order State-Estimation Problem: Given the observed system

$$\dot{x} = Ax + w_1, \quad (7.6)$$

$$y = Cx + w_2, \quad (7.7)$$

design a reduced-order state estimator

$$\dot{x}_e = A_e x_e + B_e y, \quad (7.8)$$

$$y_e = C_e x_e, \quad (7.9)$$

which minimizes the estimation criterion

$$J(A_e, B_e, C_e) \triangleq \lim_{t \rightarrow \infty} \mathbb{E}[(Lx - y_e)^T R (Lx - y_e)],$$

where $x_e \in \mathbb{R}^{n_e}$, $L \in \mathbb{R}^{p \times n}$ and L identifies the states, or linear combinations of states, whose estimates are desired. The order n_e of the estimator state x_e is determined by implementation constraints, i.e., by the computing capability available for realizing (7.8) and (7.9) in real time.

In view of the results already given, it should not be surprising (see [23]) that the optimal projection equations for reduced-order state estimation form a system of *three* matrix equations (a pair of modified Lyapunov equations along with a single modified Riccati equation) coupled by a projection which determines the gains of the optimal reduced-order estimator. This intrinsic coupling between the "operations" of optimal estimation (the modified Riccati equation) and optimal model reduction (the pair of modified Lyapunov equations) stresses the fact that reduced-order estimators designed by means of either model reduction followed by "full-order" state estimation or full-order estimation followed by estimator reduction will generally not be optimal for the given order.

VIII. DIRECTIONS FOR FURTHER RESEARCH

The most important area of research involves the further development of algorithms for solving the optimal projection

equations. Although proving local convergence of the proposed algorithm appears possible, the more important problem is achieving global optimality via the component cost approach. Although the global minimum was attained for all examples attempted by the authors, it remains to treat considerably more complex systems.

An interesting extension of the main theorem involves the case in which the original system (2.1), (2.2) is a distributed parameter system, e.g., a partial differential equation or a functional differential equation. This generalization, which has been referred to as the "ultimate reduced-order problem" [24], may lead to the efficient generation of high-order discretizations for such systems. All of the mathematical machinery required to generalize the main theorem to this case has already been applied to fixed-order dynamic compensation in [25].

IX. CONCLUSION

First-order necessary conditions for quadratically optimal reduced-order modeling of a linear time-invariant plant are expressed in the form of a pair of $n \times n$ modified Lyapunov equations coupled by an oblique projection. This form of the necessary conditions considerably simplifies the original form given by Wilson in [1] and clearly reveals the possible presence of numerous extrema. The balancing method of Moore given in [2] is shown to yield a reduced-order model that is "close" to an extremal given by the necessary conditions. A numerical example shows, however, that this extremal may very well be the global maximum rather than the desired global minimum. An algorithm is proposed which exploits the presence of the optimal projection and computes the various local extrema by the choice of eigenprojections comprising the projection. A component-cost ranking of the eigenprojections, which is very much in the spirit of Skelton's method in [3] and [12], is used to direct the algorithm to the global optimum.

It should be pointed out that Moore's balancing appears to have strong ties with the L_∞ reduction problem via the Hankel norm [29]. Alternative settings for the Hankel operator, however, seem to indicate connections to the quadratic problem [30]. Finally, the robustness problem for reduced-order modeling, estimation, and control in a quadratic setting is discussed in [31].

APPENDIX

PROOF OF THE MAIN THEOREM

Introducing the augmented system

$$\dot{\bar{x}} = \bar{A}\bar{x} + \bar{B}u,$$

$$\bar{y} = \bar{C}\bar{x},$$

where

$$\bar{x} \triangleq \begin{bmatrix} x \\ x_m \end{bmatrix}, \quad \bar{y} \triangleq y - y_m,$$

$$\bar{A} \triangleq \begin{bmatrix} A & 0 \\ 0 & A_m \end{bmatrix}, \quad \bar{B} \triangleq \begin{bmatrix} B \\ B_m \end{bmatrix}, \quad \bar{C} \triangleq [C \quad -C_m].$$

leads to the expression

$$J(A_m, B_m, C_m) = \text{tr } \bar{Q}\bar{R}, \quad (A.1)$$

where $\bar{R} \triangleq \bar{C}^T R \bar{C}$ and the nonnegative-definite steady-state covariance \bar{Q} of \bar{x} is given by the (unique) solution of

$$0 = \bar{A}\bar{Q} + \bar{Q}\bar{A}^T + \bar{V}, \quad (A.2)$$

with $\bar{V} \triangleq \bar{B}\bar{V}\bar{B}^T$. To minimize (A.1) subject to the constraint (A.2), form the Lagrangian

$$L(A_m, B_m, C_m, \bar{Q}) \triangleq \text{tr } [\lambda \bar{Q}\bar{R} + (\bar{A}\bar{Q} + \bar{Q}\bar{A}^T + \bar{V})\beta]$$

with multipliers $\lambda \geq 0$ and $\bar{P} \in \mathbb{R}^{(n+n_m) \times (n+n_m)}$. Since \mathcal{Q}_+ is an open set, the standard Lagrange multiplier rule can be applied.

Using formulas for computing partial derivatives [26], it follows that

$$0 = L_{\bar{Q}} = \bar{A}^T \bar{P} + \bar{P} \bar{A} + \lambda \bar{R}.$$

Since $\lambda = 0$ implies $\bar{P} = 0$ (recall \bar{A} is stable), we can take $\lambda = 1$ without loss of generality. Hence, \bar{P} is the (unique nonnegative-definite) solution of

$$0 = \bar{A}^T \bar{P} + \bar{P} \bar{A} + \bar{R}. \quad (\text{A.3})$$

Again using formulas from [26] and performing some manipulation, it follows that

$$0 = L_{A_m} = Q_{12}^T P_{12} + Q_2 P_2, \quad (\text{A.4})$$

$$0 = L_{B_m} = 2(P_{12}^T B + P_2 B_m) V, \quad (\text{A.5})$$

$$0 = L_{C_m} = 2R(C_m Q_2 - C Q_{12}), \quad (\text{A.6})$$

where \bar{Q} and \bar{P} have been partitioned as

$$\bar{Q} = \begin{bmatrix} Q_1 & Q_{12} \\ Q_{12}^T & Q_2 \end{bmatrix}, \quad \bar{P} = \begin{bmatrix} P_1 & P_{12} \\ P_{12}^T & P_2 \end{bmatrix}. \quad (\text{A.7})$$

Since (as will be seen shortly) Q_2 and P_2 are positive definite, define

$$G \triangleq Q_2^{-1} Q_{12}^T, \quad \Gamma \triangleq -P_2^{-1} P_{12}^T, \quad (\text{A.8})$$

so that (A.4)–(A.6) become (2.6), (2.10) and (2.11), respectively.

Next, define the nonnegative-definite matrices

$$\hat{Q} \triangleq Q_{12} Q_2^{-1} Q_{12}^T, \quad \hat{P} \triangleq P_{12} P^{-1} P_{12}^T \quad (\text{A.9})$$

and note that (A.4) implies that (2.5) holds with $M \triangleq Q_2 P_2$. Since $Q_2 P_2 = P_2^{-1/2} (P_2^{1/2} Q_2 P_2^{1/2}) P_2^{1/2}$, M is positive semisimple. The rank conditions (2.12) follow from Sylvester's inequality. Expanding (A.2) and (A.3) yields

$$0 = A Q_1 + Q_1 A^T + B V B^T, \quad (\text{A.10})$$

$$0 = A Q_{12} + Q_{12} A_m^T + B V B_m^T, \quad (\text{A.11})$$

$$0 = A_m Q_2 + Q_2 A_m^T + B_m V B_m^T, \quad (\text{A.12})$$

$$0 = A^T P_1 + P_1 A + C^T R C, \quad (\text{A.13})$$

$$0 = A^T P_{12} + P_{12} A_m - C^T R C_m, \quad (\text{A.14})$$

$$0 = A_m^T P_2 + P_2 A_m + C_m^T R C_m. \quad (\text{A.15})$$

Since A_m is stable and (A_m, B_m) is controllable, standard results (e.g., [27, p. 277]) imply that Q_2 is positive definite. Similarly, P_2 is positive definite.

It is easy to see at this point that A_m , B_m , and C_m are independent of Q_1 and P_1 and thus (A.10) and (A.13) can be ignored. Now, substituting (2.10), (2.11) and the identities

$$Q_{12} = \hat{Q} \Gamma^T, \quad P_{12} = -\hat{P} G^T, \quad (\text{A.16})$$

$$Q_2 = \Gamma \hat{Q} \Gamma^T, \quad P_2 = G \hat{P} G^T, \quad (\text{A.17})$$

into (A.11), (A.12), (A.14), and (A.15) yields

$$0 = A \hat{Q} \Gamma^T + \hat{Q} \Gamma^T A_m^T + B V B^T \Gamma^T, \quad (\text{A.18})$$

$$0 = A_m \Gamma \hat{Q} \Gamma^T + \Gamma \hat{Q} \Gamma^T A_m^T + \Gamma B V B_m^T \Gamma^T, \quad (\text{A.19})$$

$$0 = A^T \hat{P} G^T + \hat{P} G^T A_m + C^T R C G^T, \quad (\text{A.20})$$

$$0 = A_m^T \hat{P} G^T + G \hat{P} G^T A_m + G C_m^T R C G^T. \quad (\text{A.21})$$

Computing (A.19)– Γ (A.18) implies

$$A_m = \Gamma A \hat{Q} \Gamma^T (\Gamma \hat{Q} \Gamma^T)^{-1}$$

which, since $\Gamma \hat{Q} \Gamma^T = Q_2$, yields (2.9). Alternatively, (2.9) can be obtained from (A.21)– G (A.20).

If we now substitute (2.9) into (A.18)–(A.21) and use the easily verified relations (2.20), it follows that (A.19) = Γ (A.18) and (A.22) = G (A.21), and thus (A.19) and (A.21), are redundant. Finally, G^T (A.18) T and (A.20) Γ yield (2.13) and (2.14), respectively. Note that these last multiplications entail no loss of generality since $\rho(G) = \rho(\Gamma) = n_m$.

To show that the optimal projection equations entail no loss of generality over (A.2)–(A.6), let \hat{Q} , \hat{P} be extremal and define Q_{12} , Q_2 , P_{12} , P_2 by (A.16) and (A.17) for some (G, M, Γ) -factorization of $\hat{Q} \hat{P}$, and let Q_1 , P_1 satisfy (A.10) and (A.13). Then it is straightforward to reverse the steps taken in the proof to arrive at (A.2)–(A.6).

Proof of Proposition 2.1: Extremal \hat{Q} , \hat{P} leads to \hat{Q} , \hat{P} as in (A.7) satisfying (A.2)–(A.6). Computing

$$J(A_m, B_m, C_m) = \text{tr}(Q_1 C^T R C - 2 Q_{12} C_m^T R C) + \text{tr}(Q_2 C_m^T R C_m) \\ = \text{tr}[C^T R C (W_c - \hat{Q})],$$

noting that (2.13), (2.14) are equivalent to (2.21), (2.22) because of (2.20) and using (2.23), leads to (2.15). ■

REFERENCES

- [1] D. A. Wilson, "Optimum solution of model-reduction problem," *Proc. IEE*, vol. 117, pp. 1161–1165, 1970.
- [2] B. C. Moore, "Principal component analysis in linear systems: Controllability, observability, and model reduction," *IEEE Trans. Automat. Contr.*, vol. AC-26, pp. 17–32, 1981.
- [3] R. E. Skelton, "Cost Decomposition of linear systems with application to model reduction," *Int. J. Contr.*, vol. 32, pp. 1031–1055, 1980.
- [4] M. Aoki, "Control of large-scale dynamic systems by aggregation," *IEEE Trans. Automat. Contr.*, vol. AC-13, pp. 246–253, 1968.
- [5] D. A. Wilson, "Model reduction for multivariable systems," *Int. J. Contr.*, vol. 20, pp. 57–64, 1974.
- [6] R. N. Mishra and D. A. Wilson, "A new algorithm for optimal reduction of multivariable systems," *Int. J. Contr.*, vol. 31, pp. 443–466, 1980.
- [7] L. Pernebo and L. M. Silverman, "Model reduction via balanced state space representations," *IEEE Trans. Automat. Contr.*, vol. AC-27, pp. 382–387, 1982.
- [8] K. V. Fernando and H. Nicholson, "On the structure of balanced and other principal representations of SISO systems," *IEEE Trans. Automat. Contr.*, vol. AC-28, pp. 228–231, 1983.
- [9] S. Shokohi, L. M. Silverman, and P. M. Van Dooren, "Linear time-variable systems: Balancing and model reduction," *IEEE Trans. Automat. Contr.*, vol. AC-28, pp. 810–822, 1983.
- [10] E. I. Verriest and T. Kailath, "On generalized balanced realizations," *IEEE Trans. Automat. Contr.*, vol. AC-28, pp. 833–844, 1983.
- [11] E. A. Jonckheere and L. M. Silverman, "A new set of invariants for linear systems—Application to reduced-order compensator design," *IEEE Trans. Automat. Contr.*, vol. AC-28, pp. 953–964, 1983.
- [12] R. E. Skelton and A. Yousuff, "Component cost analysis of large scale systems," *Int. J. Contr.*, vol. 37, pp. 285–304, 1983.
- [13] P. T. Kabamba, "Balanced gains and their significance for balanced model reduction," in *Proc. Conf. Inform. Sci. Syst.*, Princeton Univ., Princeton, NJ, 1984; also in *IEEE Trans. Automat. Contr.*, vol. AC-30, pp. 690–693, 1985.
- [14] C. R. Rao and S. K. Mitra, *Generalized Inverse of Matrices and Its Applications*. New York: Wiley, 1971.
- [15] S. L. Campbell and C. D. Meyer, Jr., *Generalized Inverses of Linear Transformations*. London, England: Pitman, 1979.
- [16] T. Kato, *Perturbation Theory for Linear Operators*. New York: Springer-Verlag, 1966.
- [17] C. T. Mullis and R. A. Roberts, "Synthesis of minimum roundoff noise fixed point digital filters," *IEEE Trans. Circ. Syst.*, vol. CAS-23, pp. 551–562, 1976.
- [18] D. C. Hyland, "Optimality conditions for fixed-order dynamic compensation of flexible spacecraft with uncertain parameters," AIAA 20th Aerospace Sci. Meet., Orlando, FL, paper 82-0312, Jan. 1982.
- [19] —, "The optimal projection approach to fixed-order compensation: Numerical methods and illustrative results," AIAA 21st Aerospace Sci. Meet., Reno, NV, paper 83-0303, Jan. 1983.
- [20] D. C. Hyland and D. S. Bernstein, "Explicit optimality conditions for fixed-order dynamic compensation," in *Proc. IEEE Conf. Dec. Contr.*, San Antonio, TX, Dec. 1983, pp. 161–165.
- [21] —, "The optimal projection equations for fixed-order dynamic

- compensation," *IEEE Trans. Automat. Contr.*, vol. AC-29, pp. 1034-1037, 1984.
- [22] D. C. Hyland, "Comparison of various controller-reduction methods: Suboptimal versus optimal projection," in *Proc. AIAA Dynamics Specialists Conf.*, Palm Springs, CA, May 1984, pp. 381-389.
 - [23] D. S. Bernstein and D. C. Hyland, "The optimal projection equations for reduced-order state estimation," *IEEE Trans. Automat. Contr.*, vol. AC-30, pp. 583-585, 1985.
 - [24] C. S. Sims, "Reduced-order modeling and filtering," *Contr. Dynamic Syst.*, vol. 18, pp. 55-103, 1982.
 - [25] D. S. Bernstein and D. C. Hyland, "The optimal projection equations for fixed-order dynamic compensation of infinite-dimensional systems," in *Proc. AIAA Dynamics Specialists Conf.*, Palm Springs, CA, May 1984, pp. 396-400.
 - [26] M. Athans, "The matrix minimum principle," *Inform. Contr.*, vol. 11, pp. 592-606, 1968.
 - [27] W. M. Wonham, *Linear Multivariable Control: A Geometric Approach*. New York: Springer-Verlag, 1974.
 - [28] D. A. Wilson, "Reduction of the order of linear systems," Ph.D. dissertation, Univ. Cambridge, Cambridge, England, 1970.
 - [29] K. Glover, "All optimal Hankel-Norm approximations of linear multivariable systems and their L^{∞} -Error bounds," *Int. J. Contr.*, vol. 39, pp. 1115-1193, 1984.
 - [30] D. A. Wilson, "The Hankel operator and its induced norms," *Int. J. Contr.*, to be published.
 - [31] D. S. Bernstein and D. C. Hyland, "Optimal projection/maximum

entropy stochastic modelling and reduced-order design synthesis," in *Proc. IFAC Workshop Model Errors, Compensation*, Boston, MA, June 1985.



David C. Hyland received the B.S., M.S., and Sc.D. degrees in aeronautics from the Massachusetts Institute of Technology (M.I.T.), Cambridge, in 1969, 1971, 1973, respectively.

After serving as a vibration specialist in a Cambridge-based Acoustics consulting firm, he joined the M.I.T. Lincoln Laboratory Staff in 1974. His work at Lincoln Laboratory included reentry vehicle dynamics, multibody spacecraft dynamics simulation, and spacecraft attitude control. He joined Harris Corporation, Government Aerospace Systems Division in 1983, where he leads the control systems analysis and synthesis group. His current research interests include stochastic modeling of parametric uncertainty in mechanical systems and optimization theory for fixed-order dynamic compensation.

Dennis S. Bernstein (M'82), for a photograph and biography, see p. 376 of the April 1985 issue of this TRANSACTIONS.

The Optimal Projection Equations for Reduced-Order State Estimation

DENNIS S. BERNSTEIN AND DAVID C. HYLAND

Abstract—First-order necessary conditions for optimal, steady-state, reduced-order state estimation for a linear, time-invariant plant in the presence of correlated disturbance and nonsingular measurement noise are derived in a new and highly simplified form. In contrast to the lone matrix Riccati equation arising in the full-order (Kalman filter) case, the optimal steady-state reduced-order estimator is characterized by three matrix equations (one modified Riccati equation and two modified Lyapunov equations) coupled by a projection whose rank is precisely equal to the order of the estimator and which determines the optimal estimator gains. This coupling is a graphic reminder of the suboptimality of proposed approaches involving either model reduction followed by "full-order" estimator design or full-order estimator design followed by estimator-reduction techniques. The results given here complement recently obtained results which characterize the optimal reduced-order model by means of a pair of coupled modified Lyapunov equations [7] and the optimal fixed-order dynamic compensator by means of a coupled system of two modified Riccati equations and two modified Lyapunov equations [6].

I. INTRODUCTION

It has recently been shown (see [1]–[7]) that the first-order necessary conditions for the problems of optimal model reduction and optimal fixed-order dynamic compensation can be formulated in terms of an "optimal

Manuscript received June 28, 1984; revised September 14, 1984. This work was supported in part by Lincoln Laboratory, M.I.T.

The authors are with Harris Corporation, GASD, Melbourne, FL 32901.

projection" matrix which arises as a direct consequence of optimality. These necessary conditions, by virtue of their remarkable simplicity, yield insight into the structure of the optimal design and permit the development of alternative numerical algorithms [2], [4], [7]. The purpose of this note is to develop analogous first-order necessary conditions for the reduced-order state-estimation problem. Since this problem falls midway between the problems of open-loop model reduction and closed-loop fixed-order dynamic compensation, it is not surprising that the necessary conditions for these problems are correspondingly related. Specifically, while the optimal projection equations for model reduction consist of a system of two matrix equations (a pair of modified Lyapunov equations) and the optimal projection equations for fixed-order dynamic compensation comprise a system of four matrix equations (a pair of modified Lyapunov equations plus a pair of modified Riccati equations), the optimal projection equations for reduced-order state estimation form a system of three matrix equations (a pair of modified Lyapunov equations along with a single modified Riccati equation). In each case the system of matrix equations is coupled by an oblique projection (idempotent matrix) which determines the gains of the optimal reduced-order system, whether it be a model, estimator, or compensator.

The need for designing an optimal reduced-order state estimator for a high-order dynamic system follows directly from real-world constraints on computing capability. A further motivation is the fact that although a system may have many degrees of freedom, it is often the case that estimates of only a small number of state variables are actually required. In the face of these practical motivations, numerous approaches to designing reduced-order state estimators have been proposed. See [8] for a recent review of previous results.

An important fact pointed out in [8] and [9] is that reduced-order estimators designed by means of either model reduction followed by "full-order" state estimation or full-order estimation followed by estimator reduction will not be optimal for the given order. In the present paper this point is graphically confirmed by the fact that the three matrix equations characterizing the optimal reduced-order state estimator reveal intrinsic coupling (via the optimal projection) between the "operations" of optimal estimation (the modified Riccati equation) and optimal model reduction (the pair of modified Lyapunov equations).

II. PROBLEM STATEMENT AND MAIN RESULT

The following notation and definitions will be used throughout the paper:

| | |
|---|--|
| n, l, n_e, p | positive integers, $1 \leq n_e \leq n$ |
| x, y, x_e, y_e | n, l, n_e, p -dimensional vectors |
| A, C, L | $n \times n, l \times n, p \times n$ real matrices |
| A_e, B_e, C_e | $n_e \times n_e, n_e \times l, p \times n_e$ real matrices |
| $w_1(t), t \geq 0$ | n -dimensional white noise with nonnegative-definite intensity V_1 |
| $w_2(t), t \geq 0$ | l -dimensional white noise with positive-definite intensity V_2 |
| V_{12} | $n \times l$ matrix satisfying $E[w_1(t)w_2(s)^T] = V_{12}\delta(t-s)$ |
| R | $p \times p$ positive-definite matrix |
| I, I_r | $r \times r$ identity matrix |
| Z^T | transpose of vector or matrix Z |
| Z^{-T} | $(Z^T)^{-1}$ or $(Z^{-1})^T$ |
| $\mathcal{R}(Z), \mathcal{R}(Z), \rho(Z)$ | null space, range, rank of matrix Z |
| \mathbb{E} | expected value |
| $\mathbb{R}, \mathbb{R}^{r \times s}$ | real numbers, $r \times s$ real matrices |
| stable matrix | matrix with eigenvalues in open left half plane |
| nonnegative-definite matrix | symmetric matrix with nonnegative eigenvalues |
| positive-definite matrix | symmetric matrix with positive eigenvalues |
| nonnegative-semisimple matrix | matrix similar to a nonnegative-definite matrix |

positive-semisimple matrix matrix similar to a positive-definite matrix
positive-diagonal matrix diagonal matrix with positive diagonal elements

We consider the following optimal reduced-order state-estimation problem. Given the system

$$\dot{x} = Ax + w_1, \quad (2.1)$$

$$y = Cx + w_2, \quad (2.2)$$

design a reduced-order state estimator

$$\dot{x}_e = A_e x_e + B_e y, \quad (2.3)$$

$$y_e = C_e x_e, \quad (2.4)$$

which minimizes the error criterion

$$J(A_e, B_e, C_e) \triangleq \lim_{t \rightarrow \infty} \mathbb{E}[(Lx - y_e)^T R (Lx - y_e)].$$

In this formulation the matrix L identifies the states, or linear combinations of states, whose estimates are desired. The order n_e of the estimator state x_e is determined by implementation constraints, i.e., by the computing capability available for realizing (2.3), (2.4) in real time. Hence, n_e is considered to be fixed in what follows and the problem is concerned with determining A_e , B_e , and C_e .

To guarantee that J is finite it is assumed that A is stable and we restrict our attention to the set of stable reduced-order estimators

$$\mathcal{Q} \triangleq \{(A_e, B_e, C_e) : A_e \text{ is stable}\}.$$

Since the value of J is independent of the internal realization of the transfer function corresponding to (2.3) and (2.4), without loss of generality we further restrict our attention to the set of admissible estimators

$$\mathcal{C} \triangleq \{(A_e, B_e, C_e) \in \mathcal{Q} :$$

$$(A_e, B_e) \text{ is controllable and } (A_e, C_e) \text{ is observable}\}.$$

The following lemma, whose proof is given in [7], is needed for the statement of the main result.

Lemma 2.1: Suppose $\hat{Q}, \hat{P} \in \mathbb{R}^{n \times n}$ are nonnegative definite. Then $\hat{Q}\hat{P}$ is nonnegative semisimple. Furthermore, if $\rho(\hat{Q}\hat{P}) = n_e$, then there exist $G, \Gamma \in \mathbb{R}^{n_e \times n}$ and positive-semisimple $M \in \mathbb{R}^{n_e \times n_e}$ such that

$$\hat{Q}\hat{P} = G^T M \Gamma, \quad (2.5)$$

$$\Gamma G^T = I_{n_e}. \quad (2.6)$$

For convenience in stating the Main Theorem we shall refer to $G, \Gamma \in \mathbb{R}^{n_e \times n}$ and positive-semisimple $M \in \mathbb{R}^{n_e \times n_e}$ satisfying (2.5) and (2.6) as a (G, M, Γ) -factorization of $\hat{Q}\hat{P}$. Furthermore, define the notation

$$\tau \triangleq G^T \Gamma, \quad \tau_1 \triangleq I_n - \tau$$

and

$$Q \triangleq Q C^T + V_{12},$$

where $Q \in \mathbb{R}^{n \times n}$.

Main Theorem: Suppose $(A_e, B_e, C_e) \in \mathcal{C}$, solves the optimal reduced-order state-estimation problem. Then there exist nonnegative-definite matrices $\hat{Q}, \hat{P} \in \mathbb{R}^{n_e \times n_e}$ such that, for some (G, M, Γ) -factorization of $\hat{Q}\hat{P}$, A_e, B_e , and C_e are given by

$$A_e = \Gamma(A - QV_2^{-1}C)G^T, \quad (2.7)$$

$$B_e = \Gamma QV_2^{-1}, \quad (2.8)$$

$$C_e = LG^T \quad (2.9)$$

and such that the following conditions are satisfied:

$$0 = AQ + QA^T + V_1 - QV_2^{-1}Q^T + \tau_1 QV_2^{-1}Q^T \tau_1^T, \quad (2.10)$$

$$0 = A\hat{Q} + \hat{Q}A^T + QV_2^{-1}Q^T - \tau_1 QV_2^{-1}Q^T \tau_1^T, \quad (2.11)$$

$$0 = (A - QV_2^{-1}C)^T \hat{P} + \hat{P}(A - QV_2^{-1}C) + L^T R L - \tau_1^T L^T R L \tau_1, \quad (2.12)$$

$$\rho(\hat{Q}) = \rho(\hat{P}) = \rho(\hat{Q}\hat{P}) = n_e. \quad (2.13)$$

Remark 2.1: It is useful to note that (2.7) can be replaced by

$$A_e = \Gamma A G^T - B_e C G^T. \quad (2.7)'$$

Remark 2.2: Because of (2.6) the $n \times n$ matrix τ which couples the three equations (2.10)–(2.12) is idempotent, i.e., $\tau^2 = \tau$. In general, this "optimal projection" is an oblique projection (as opposed to an orthogonal projection) since it is not necessarily symmetric. Note that from Sylvester's inequality and (2.6) it follows that $\rho(\tau) = n_e$. It should be stressed that the form of the optimal reduced-order estimator (2.7)–(2.9) is a direct consequence of optimality and not the result of an *a priori* assumption on the structure of the reduced-order estimator.

Remark 2.3: To obtain the standard steady-state Kalman filter result for the full-order case, set $p = n_e = n$ and $L = I_n$. Then $\tau = G = \Gamma = I_n$ and thus (2.10) reduces to the standard observer Riccati equation [10, p. 367] and (2.7) and (2.8) yield the usual expressions. Furthermore, it follows from (2.7)' [11, Lemma 2.1] and standard results that (2.11)–(2.13) are equivalent to the assumption that (A_e, B_e, C_e) is controllable and observable.

Remark 2.4: Since $\hat{Q}\hat{P}$ is nonnegative semisimple it has a group generalized inverse $(\hat{Q}\hat{P})^\#$ given by $G^T M^{-1} \Gamma$ (see, e.g., [12, p. 124]). Hence, by (2.6) the optimal projection τ is given by

$$\tau = \hat{Q}\hat{P}(\hat{Q}\hat{P})^\#. \quad (2.14)$$

Remark 2.5: Replacing x_e by Sx_e , where S is invertible, yields the "equivalent" estimator $(SA_e S^{-1}, SB_e, C_e S^{-1})$. Since $J(A_e, B_e, C_e) = J(SA_e S^{-1}, SB_e, C_e S^{-1})$, one would expect the Main Theorem to apply also to $(SA_e S^{-1}, SB_e, C_e S^{-1})$. This is indeed the case since transformation of the estimator state basis corresponds to the alternative factorization $\hat{Q}\hat{P} = (S^{-T}G)^T (SMS^{-1})(ST)$.

Remark 2.6: Note that, for the optimal values of A_e, B_e , and C_e , (2.3) assumes the observer form

$$\dot{x}_e = \Gamma A G^T x_e + \Gamma QV_2^{-1}(y - C G^T x_e). \quad (2.15)$$

By introducing the quasi-full-state estimate $\hat{x} \triangleq G^T x_e \in \mathbb{R}^{n_e}$ so that $\tau \hat{x} = \hat{x}$ and $x_e = \Gamma \hat{x} \in \mathbb{R}^{n_e}$, (2.15) can be written as

$$\dot{\hat{x}} = \tau A \tau \hat{x} + \tau QV_2^{-1}(y - C \hat{x}). \quad (2.16)$$

Note that although the implemented estimator (2.15) has the state $x_e \in \mathbb{R}^{n_e}$, (2.15) can be viewed as a quasi-full-order estimator whose geometric structure is entirely dictated by the projection τ . Specifically, error inputs $QV_2^{-1}(y - C \hat{x})$ are annihilated unless they are contained in $[\mathcal{N}(\tau)]^\perp = \mathcal{R}(\tau^T)$. Hence, the observation subspace of the estimator is precisely $\mathcal{R}(\tau^T)$.

Remark 2.7: Although the form of (2.16) would lead one to surmise that the optimal reduced-order estimator is a projection of the optimal full-order estimator, this is not generally the case for the following simple reason. In the full-order case Q (which appears in \mathcal{Q}) is determined by solving a single Riccati equation, whereas in the reduced-order case Q must be found in conjunction with \hat{Q} and \hat{P} to satisfy all three matrix equations (2.10)–(2.12). Hence, the value of Q in the reduced-order case may be different from the value of Q in the full-order case. Thus, (2.16) may not be obtainable by simply projecting the full-order result.

To further clarify the relationship between \hat{Q}, \hat{P} , and τ , we now show that there exists a similarity transformation which simultaneously diagonalizes $\hat{Q}\hat{P}$ and τ .

Proposition 2.1: There exists invertible $\Phi \in \mathbb{R}^{n_e \times n_e}$ such that

$$\hat{Q} = \Phi^{-1} \begin{bmatrix} \Lambda & 0 \\ 0 & 0 \end{bmatrix} \Phi^{-T}, \quad \hat{P} = \Phi^T \begin{bmatrix} \Lambda & 0 \\ 0 & 0 \end{bmatrix} \Phi, \quad (2.17)$$

$$\hat{Q}\hat{P} = \Phi^{-1} \begin{bmatrix} \Lambda & 0 \\ 0 & 0 \end{bmatrix} \Phi, \quad \tau = \Phi^{-1} \begin{bmatrix} I_{n_e} & 0 \\ 0 & 0 \end{bmatrix} \Phi. \quad (2.18a,b)$$

where $\Lambda_Q, \Lambda_P \in \mathbb{R}^{n \times n}$ are positive diagonal, $\Lambda \triangleq \Lambda_Q \Lambda_P$, and the diagonal elements of Λ are the eigenvalues of M . Consequently,

$$\dot{Q} = \tau Q, \quad \dot{P} = P\tau. \quad (2.19)$$

III. PROOF OF THE MAIN THEOREM

The proof proceeds exactly as in [6]. Using the fact that Q_+ is open, the Fritz John version of the Lagrange multiplier theorem can be used to rigorously derive the first-order necessary conditions

$$0 = \bar{A}\bar{Q} + \bar{Q}\bar{A}^T + \bar{V}, \quad (3.1)$$

$$0 = \bar{A}^T \bar{P} + \bar{P} \bar{A} + \bar{R}, \quad (3.2)$$

$$0 = P_{12}^T Q_{12} + P_2 Q_2, \quad (3.3)$$

$$B_r = -(P_2^{-1} P_{12}^T Q_1 + Q_2^T C^T + P_2^{-1} P_{12}^T V_{12}) V_2^{-1}, \quad (3.4)$$

$$C_r = L Q_{12} Q_2^{-1}, \quad (3.5)$$

where

$$\bar{A} = \begin{bmatrix} A & 0 \\ B_r C & A_r \end{bmatrix}, \quad \bar{V} = \begin{bmatrix} V_1 & V_{12} B_r^T \\ B_r V_{12} & B_r V_2 B_r^T \end{bmatrix},$$

$$\bar{R} = \begin{bmatrix} L^T R L & -L^T R C_r \\ -C_r^T R L & C_r^T R C_r \end{bmatrix}$$

and $(n + n_r) \times (n + n_r) \bar{Q}, \bar{P}$ are partitioned into $n \times n, n \times n_r$, and $n_r \times n_r$ subblocks as

$$\bar{Q} = \begin{bmatrix} Q_1 & Q_{12} \\ Q_{12}^T & Q_2 \end{bmatrix}, \quad \bar{P} = \begin{bmatrix} P_1 & P_{12} \\ P_{12}^T & P_2 \end{bmatrix}.$$

Expanding (3.1) and (3.2) yields

$$0 = A Q_1 + Q_1 A^T + V_1, \quad (3.6)$$

$$0 = A Q_{12} + Q_{12} A_r^T + Q_1 (B_r C)^T + V_{12} B_r^T, \quad (3.7)$$

$$0 = A_r Q_2 + Q_2 A_r^T + B_r C Q_{12} + Q_{12}^T (B_r C)^T + B_r V_2 B_r^T, \quad (3.8)$$

$$0 = A^T P_1 + P_1 A + (B_r C)^T P_{12}^T + P_{12} B_r C + L^T R L, \quad (3.9)$$

$$0 = P_{12} A_r + A^T P_{12} + (B_r C)^T P_2 - L^T R C_r, \quad (3.10)$$

$$0 = A_r^T P_2 + P_2 A_r + C_r^T R C_r. \quad (3.11)$$

Note that (3.9) is superfluous and can be omitted. Writing (3.8) as (see [13], [14])

$$0 = (A_r + B_r C Q_{12} Q_2^{-1}) Q_2 + Q_2 (A_r + B_r C Q_{12} Q_2^{-1})^T + B_r V_2 B_r^T,$$

where Q_2^{-1} is the Moore-Penrose or Drazin generalized inverse of Q_2 , it follows from [11, Lemmas 2.1 and 12.2] that Q_2 is positive definite. Similarly, (3.11) implies that P_2 is positive definite. This justifies (3.4) and (3.5).

Now define the $n \times n$ nonnegative-definite matrices (see [13], [14])

$$Q = Q_1 - Q_{12} Q_2^{-1} Q_{12}^T, \quad \bar{Q} = Q_{12} Q_2^{-1} Q_{12}^T, \quad \bar{P} = P_{12} P_2^{-1} P_{12}^T,$$

and note that (3.3) implies (2.5) and (2.6) with

$$G = Q_2^{-1} Q_{12}^T, \quad M = Q_2 P_2, \quad \Gamma = -P_2^{-1} P_{12}^T.$$

Since $Q_2 P_2 = P_2^{-1/2} (P_2^{1/2} Q_2 P_2^{1/2}) P_2^{1/2}$, M is positive semisimple. Sylvester's inequality yields (2.13). Note (2.19) and the identities

$$Q_1 = Q + \bar{Q}, \quad (3.12)$$

$$Q_{12} = \bar{Q} \Gamma^T, \quad P_{12} = -\bar{P} G^T, \quad (3.13)$$

$$Q_2 = \Gamma \bar{Q} \Gamma^T, \quad P_2 = G \bar{P} G^T. \quad (3.14)$$

Using (3.12)–(3.14), (3.4) and (3.5) yield (2.8) and (2.9). Also, the right-hand sides of (3.8) and (3.7) yield (2.7). Substituting (2.7)–(2.9) into (3.6)–(3.8), (3.10) and (3.11), it can be seen that (3.8) and (3.11) are also superfluous. Finally, linear combinations of the remaining three equations (3.6), (3.7), and (3.10) yield (2.10)–(2.12).

IV. CONCLUDING REMARKS

The question of multiple local minima satisfying the optimal projection equations for reduced-order state estimation and the problem of constructing numerical methods for solving these equations are beyond the scope of this note. It should be pointed out, however, that promising numerical results for the model-reduction and fixed-order dynamic-compensation problems have been obtained by means of iterative algorithms that take full advantage of the presence and structure of the optimal projection [2], [4], [7].

Finally, the results of this paper can be extended to include the following related problems: 1) discrete-time system/discrete-time estimator; 2) infinite-dimensional system/finite-dimensional estimator [5]; and 3) parameter uncertainties [1], [15], [16].

ACKNOWLEDGMENT

The authors wish to thank Dr. F. M. Ham for directing their attention to the reduced-order state-estimation problem as a fruitful application of the optimal projection approach.

REFERENCES

- [1] D. C. Hyland, "Optimality conditions for fixed-order dynamic compensation of flexible spacecraft with uncertain parameters," *AIAA 20th Aerosp. Sci. Meet.*, Orlando, FL, Jan. 1982, paper 82-0312.
- [2] —, "The optimal projection approach to fixed-order compensation: Numerical methods and illustrative results," *AIAA 21st Aerosp. Sci. Meet.*, Reno, NV, Jan. 1983, paper 83-0303.
- [3] D. C. Hyland and D. S. Bernstein, "Explicit optimality conditions for fixed-order dynamic compensation," in *Proc. 22nd IEEE Conf. Decision Contr.*, San Antonio, TX, Dec. 1983, pp. 161–165.
- [4] D. C. Hyland, "Comparison of various controller-reduction methods: Suboptimal versus optimal projection," in *Proc. AIAA Dynam. Specialists Conf.*, Palm Springs, CA, May 1984, pp. 381–389.
- [5] D. S. Bernstein and D. C. Hyland, "The optimal projection equations for fixed-order dynamic compensation of distributed parameter systems," presented at the AIAA Dynam. Specialists Conf., Palm Springs, CA, May 1984.
- [6] D. C. Hyland and D. S. Bernstein, "The optimal projection equations for fixed-order dynamic compensation," *IEEE Trans. Automat. Contr.*, vol. AC-29, pp. 1034–1037, 1984.
- [7] —, "The optimal projection approach to model reduction and the relationship between the methods of Wilson and Moore," in *Proc. 23rd Conf. Decision Contr.*, Las Vegas, NV, Dec. 1984.
- [8] C. S. Sims, "Reduced-order modelling and filtering," in *Control and Dynamic Systems*, Vol. 18, C. T. Leondes, Ed., 1982, pp. 55–103.
- [9] D. A. Wilson and R. N. Mishra, "Design of low order estimators using reduced models," *Int. J. Contr.*, vol. 23, pp. 447–456, 1979.
- [10] K. Kwakernaak and R. Sivan, *Linear Optimal Control Systems*. New York: Wiley-Interscience, 1972.
- [11] W. M. Wonham, *Linear Multivariable Control: A Geometric Approach*. New York: Springer-Verlag, 1974.
- [12] S. L. Campbell and C. D. Meyer, Jr., *Generalized Inverses of Linear Transformations*. London: Pitman, 1979.
- [13] A. Albert, "Conditions for positive and nonnegative definiteness in terms of pseudo inverse," *SIAM J. Appl. Math.*, vol. 17, pp. 434–440, 1969.
- [14] E. Kreindler and A. Jameson, "Conditions for nonnegativeness of partitioned matrices," *IEEE Trans. Automat. Contr.*, vol. AC-17, pp. 147–148, 1972.
- [15] P. J. McLane, "Optimal linear filtering for linear systems with state-dependent noise," *Int. J. Contr.*, vol. 10, pp. 42–51, 1969.
- [16] D. S. Bernstein and D. C. Hyland, "Optimal projection/maximum entropy stochastic modeling and reduced-order design synthesis," *IFAC Workshop on Model Error and Concepts and Compensation*, Boston, MA, June 1985.

The Optimal Projection Equations for Fixed-Order Dynamic Compensation

DAVID C. HYLAND AND DENNIS S. BERNSTEIN

Abstract—First-order necessary conditions for quadratically optimal, steady-state, fixed-order dynamic compensation of a linear, time-invariant plant in the presence of disturbance and observation noise are derived in a new and highly simplified form. In contrast to the pair of matrix Riccati equations for the full-order LQG case, the optimal steady-state fixed-order dynamic compensator is characterized by four matrix equations (two modified Riccati equations and two modified Lyapunov equations) coupled by a projection whose rank is precisely equal to the order of the compensator and which determines the optimal compensator gains. The coupling represents a graphic portrayal of the demise of the classical separation principle for the reduced-order controller case.

I. INTRODUCTION

Because of constraints imposed by on-line computations, dynamic controllers for high-order systems such as flexible spacecraft must be of relatively modest order. Hence, this paper is concerned with the design of quadratically optimal, fixed-order (i.e., reduced-order) dynamic compensation for a plant subject to stochastic disturbances and nonsingular measurement noise. Since white noise in all measurement channels precludes direct output feedback (see Section II), only purely dynamic controllers are considered. The requirements for resolution of this optimization problem include the following.

1) Conditions for the existence of an optimal, stabilizing compensator of the prescribed order. (In the full-order case these are the usual stabilizability and detectability conditions of LQG theory.)

2) Stationary conditions, i.e., first-order necessary conditions, rendered in a tractable form to facilitate developments in items 3) and 4) below. (In the full-order case these conditions are precisely the LQG gain relations together with the regulator and observer Riccati equations.)

3) Sufficiency conditions, i.e., additional restrictions on solutions of the first-order necessary conditions which characterize local minima and single out the *global* minimum. (In the full-order case the global minimum is distinguished by the unique nonnegative-definite solutions to the LQG Riccati equations.)

4) Convergent numerical algorithms for simultaneous satisfaction of the necessary and sufficient conditions. (In the full-order case numerical algorithms have been devised which take full advantage of the highly structured form of the Riccati equations.)

Manuscript received August 15, 1983; revised February 14, 1983. This paper is based on a prior submission of March 10, 1983. Paper recommended by P. R. Kumar, Past Chairman of the Stochastic Control Committee. This work was supported by the Department of the Air Force and was performed at Lincoln Laboratory, M I T.

The authors are with the Controls Analysis and Synthesis Group, Harris Corp., GASD, Melbourne, FL 32902.

The present paper deals exclusively with item 2). Although the stationary conditions for the fixed-order compensation problem have been written down (see [11]–[12], for example), full exploitation has undoubtedly been impeded by their extreme complexity [see (3.3)–(3.11)]. What has been lacking, to quote the insightful remarks of [9], “is a deeper understanding of the structural coherence of these equations.” The contribution of the present paper is to show how the originally very complex stationary conditions can be transformed without loss of generality to much simpler and more tractable forms. The resulting equations (2.10)–(2.17) preserve the simple form of LQG relations for the gains in terms of covariance and cost matrices which, in turn, are determined by a coupled system of two modified Riccati equations and two modified Lyapunov equations. This coupling, by means of a projection (idempotent matrix) whose rank is precisely equal to the order of the compensator, represents a graphic portrayal of the demise of the classical separation principle for the reduced-order controller case. When, as a special case, the order of the compensator is required to be equal to the order of the plant, the modified Riccati equations reduce to the standard LQG Riccati equations and the modified Lyapunov equations express the proviso that the compensator be minimal, i.e., controllable and observable. Since the LQG Riccati equations as such are nothing more than the necessary conditions for full-order compensation, we believe that the “optimal projection equations” provide a clear and simple generalization of standard LQG theory.

Since we are concerned with optimal fixed-order compensator design, our approach does not represent yet another model- or controller-reduction scheme along the lines of [13]–[17]. Indeed, the optimal projection equations, by virtue of their relatively transparent structure, can reveal the extent to which the design equations of a given ad hoc reduction scheme conform to the necessary conditions for optimality. For example, the oblique projection which arises in the present formulation may not be of the form $\begin{bmatrix} I & 0 \\ 0 & 0 \end{bmatrix}$ even in the basis corresponding to the “balanced” realization [13]–[16]. These issues are discussed in [18] where the results of [19] are simplified by means of the approach of the present paper and where the balancing method of [13] is reinterpreted in the context of optimality theory.

The fact that the optimal projection equations consist of four coupled matrix equations, i.e., two modified Riccati equations and two modified Lyapunov equations, should not be at all surprising for the following simple reason. Reduced-order control-design methods often involve either LQG applied to a reduced-order model or model reduction applied to a full-order LQG design. Both approaches, then, involve the solution of precisely four equations: two Riccati equations (for LQG) plus two Lyapunov equations (for model reduction via balancing, as in [13]). The coupled form of the optimal projection equations is thus a strong reminder that the LQG and order-reduction operations cannot be iterated but must, in a certain sense, be performed simultaneously.

II. PROBLEM STATEMENT AND THE MAIN THEOREM

Given the control system

$$\dot{x}(t) = Ax(t) + Bu(t) + w_1(t), \quad (2.1)$$

$$y(t) = Cx(t) + w_2(t) \quad (2.2)$$

design a fixed-order dynamic compensator

$$\dot{x}_c(t) = A_c x_c(t) + B_c y(t), \quad (2.3)$$

$$u(t) = C_c x_c(t) \quad (2.4)$$

which minimizes the steady-state performance criterion

$$J(A_c, B_c, C_c) \triangleq \lim_{t \rightarrow \infty} \mathbb{E} [x(t)^T R_1 x(t) + u(t)^T R_2 u(t)] \quad (2.5)$$

where: $x \in \mathbb{R}^n$, $u \in \mathbb{R}^m$, $y \in \mathbb{R}^l$, $x_c \in \mathbb{R}^{n_c}$, $n_c \leq n$, $A, B, C, A_c, B_c, C_c, R_1$, and R_2 are matrices of appropriate dimension with R_1 (symmetric) nonnegative definite and R_2 (symmetric) positive definite; w_1 is white disturbance noise with $n \times n$ nonnegative-definite intensity V_1 and w_2 is

white observation noise with $l \times l$ positive-definite intensity V_2 ; w_1 and w_2 are uncorrelated and have zero mean. We note that the assumptions of nonsingular control weighting and nonsingular observation noise preclude the use of direct output feedback as in

$$u(t) = C_c x_c(t) + D_c y(t) \quad (2.6)$$

since J is undefined unless (see [7])

$$\text{tr} [D_c^T R_2 D_c V_2] = 0 \quad (=\ R_2 D_c V_2 = 0). \quad (2.7)$$

To guarantee that J is finite and independent of initial conditions we restrict our attention to the set of admissible stabilizing compensators

$$\mathcal{A} \triangleq \left\{ (A_c, B_c, C_c) : \tilde{A} \triangleq \begin{bmatrix} A & B C_c \\ B_c C & A_c \end{bmatrix} \text{ is asymptotically stable} \right\}$$

where \tilde{A} is the closed-loop dynamics matrix. Since the value of J is independent of the internal realization of the compensator, we can further restrict our attention to

$$\mathcal{A}_+ \triangleq \{ (A_c, B_c, C_c) \in \mathcal{A} :$$

$$(A_c, B_c) \text{ is controllable and } (C_c, A_c) \text{ is observable} \}.$$

For the following lemma call a square matrix nonnegative (respectively, positive) semisimple if it has a diagonal Jordan form and nonnegative (respectively, positive) eigenvalues. Let I_r denote the $r \times r$ identity matrix.

Lemma 2.1: Suppose $\hat{Q}, \hat{P} \in \mathbb{R}^{n \times n}$ are nonnegative definite. Then $\hat{Q}\hat{P}$ is nonnegative semisimple. Furthermore, if $\text{rank } \hat{Q}\hat{P} = n_c$ then there exist $G, \Gamma \in \mathbb{R}^{n_c \times n}$ and positive-semisimple $M \in \mathbb{R}^{n_c \times n_c}$ such that

$$\hat{Q}\hat{P} = G^T M \Gamma, \quad (2.8)$$

$$\Gamma G^T = I_{n_c}. \quad (2.9)$$

Proof. The result is an immediate consequence of [20, Theorem 6.2.5, p. 123]. \square

For convenience in stating the Main Theorem, define

$$\Sigma \triangleq B R_2^{-1} B^T, \quad \Omega \triangleq C^T V_2^{-1} C$$

and call G, M , and Γ satisfying (2.8) and (2.9) a (G, M, Γ) -factorization of $\hat{Q}\hat{P}$.

Main Theorem: Suppose $(A_c, B_c, C_c) \in \mathcal{A}_+$ solves the steady-state fixed-order dynamic-compensation problem. Then there exist $n \times n$ nonnegative-definite matrices Q, P, \hat{Q} , and \hat{P} such that A_c, B_c , and C_c are given by

$$A_c = \Gamma(A - Q\Sigma - \Sigma P)G^T, \quad (2.10)$$

$$B_c = \Gamma Q C^T V_2^{-1}, \quad (2.11)$$

$$C_c = -R_2^{-1} B^T P G^T \quad (2.12)$$

for some (G, M, Γ) -factorization of $\hat{Q}\hat{P}$, and such that with $\tau \triangleq G^T \Gamma$ the following conditions are satisfied:

$$0 = (A - \tau Q \Sigma) Q + Q(A - \tau Q \Sigma)^T + V_1 + \tau Q \Sigma Q^T \tau^T, \quad (2.13)$$

$$0 = (A - \Sigma P \tau)^T P + P(A - \Sigma P \tau) + R_1 + \tau^T P \Sigma P \tau, \quad (2.14)$$

$$0 = \tau[(A - \Sigma P)\hat{Q} + \hat{Q}(A - \Sigma P)^T + Q \Sigma Q], \quad (2.15)$$

$$0 = [(A - Q \Sigma)^T \hat{P} + \hat{P}(A - Q \Sigma) + P \Sigma P] \tau, \quad (2.16)$$

$$\text{rank } \hat{Q} = \text{rank } \hat{P} = \text{rank } \hat{Q}\hat{P} = n_c. \quad (2.17)$$

Remark 2.1: Because of (2.9) the $n \times n$ matrix τ which couples the four equations (2.13)–(2.16) is idempotent, i.e., $\tau^2 = \tau$. In general this “optimal projection” is an oblique projection (as opposed to an orthogonal projection) since it is not necessarily symmetric. Note that Sylvester’s inequality and (2.9) imply that $\text{rank } \tau = n_c$.

Remark 2.2: Using the relations $\hat{Q} = \tau \hat{Q}$ and $\hat{P} = \hat{P} \tau$ [see (3.12)],

the optimal projection equations (2.13)–(2.16) can be written in the equivalent form

$$0 = A Q + Q A^T + V_1 - Q \Sigma Q + \tau_1 Q \Sigma Q \tau_1^T, \quad (2.18)$$

$$0 = A^T P + P A + R_1 - P \Sigma P + \tau_1^T P \Sigma P \tau_1, \quad (2.19)$$

$$0 = (A - \Sigma P) \hat{Q} + \hat{Q} (A - \Sigma P)^T + Q \Sigma Q - \tau_1 Q \Sigma Q \tau_1^T, \quad (2.20)$$

$$0 = (A - Q \Sigma)^T \hat{P} + \hat{P} (A - Q \Sigma) + P \Sigma P - \tau_1^T P \Sigma P \tau_1, \quad (2.21)$$

where $\tau_1 \triangleq I_n - \tau$. Note that in the full-order case $n_c = n$, $\tau = G = \Gamma = I_n$ and thus (2.18) and (2.19) reduce to the standard observer and regulator Riccati equations and (2.10)–(2.12) yield the usual LQG expressions. Furthermore, it can be shown that (2.20), (2.21), and (2.17) are equivalent to the assumption that (A_c, B_c, C_c) is controllable and observable.

Remark 2.3: Since $\hat{Q}\hat{P}$ is nonnegative semisimple it has a group generalized inverse $(\hat{Q}\hat{P})^\#$ given by $G^T M^{-1} \Gamma$ (see e.g., [21, p. 124]). Hence, by (2.9) the optimal projection τ is given by

$$\tau = \hat{Q}\hat{P}(\hat{Q}\hat{P})^\#. \quad (2.22)$$

Remark 2.4: The modified Riccati equations (2.13) and (2.14) are similar to the (single) "extended algebraic Riccati equation" which arises in the static output feedback problem (see, e.g., [22]).

Remark 2.5: Replacing x_c by Sx_c , where S is invertible, yields the "equivalent" compensator $(SA_c S^{-1}, SB_c, C_c S^{-1})$. Since $J(A_c, B_c, C_c) = J(SA_c S^{-1}, SB_c, C_c S^{-1})$ one would expect the Main Theorem to apply also to $(SA_c S^{-1}, SB_c, C_c S^{-1})$. This is indeed the case since transformation of the compensator state basis corresponds to the alternative factorization $\hat{Q}\hat{P} = (S^{-T}G)^T (SMS^{-1}) (ST)$. See [10] for related remarks.

Remark 2.6: By introducing the quasi-full-state estimate $\hat{x} \triangleq G^T x_c \in \mathbb{R}^n$ so that $\tau \hat{x} = \hat{x}$ and $x_c = \Gamma \hat{x} \in \mathbb{R}^{n_c}$, (2.1)–(2.4) can be written as

$$\dot{\hat{x}} = A \hat{x} + B C_c \tau \hat{x} + w_1,$$

$$\hat{x} = \tau(A - \hat{B}_c C + B C_c) \tau \hat{x} + \tau \hat{B}_c (C x + w_2)$$

where $\hat{B}_c \triangleq Q C^T V_2^{-1}$ and $\hat{C}_c \triangleq -R_2^{-1} B^T P$. Although the implemented compensator has the state $x_c \in \mathbb{R}^{n_c}$, it can be viewed as a quasi-full-order compensator whose geometric structure is entirely dictated by the projection τ . Sensor inputs $\hat{B}_c y$ are annihilated unless they are contained in $[\mathcal{N}(\tau)]^\perp = \mathcal{R}(\tau^T)$, where \mathcal{N} and \mathcal{R} denote null space and range. Furthermore, the quasi-full-order state estimate $\tau \hat{x}$ employed in the control input is contained in $\mathcal{R}(\tau)$. Thus, $\mathcal{R}(\tau)$ and $\mathcal{R}(\tau^T)$ are the control and observation subspaces of the compensator.

III. PROOF OF THE MAIN THEOREM

The proof given here considerably simplifies the original derivation given in [23] and [24]. Using the fact that \mathcal{Q}_2 is open, the Fritz John version of the Lagrange multiplier theorem can be used to rigorously derive the first-order necessary conditions ([7], see also [25])

$$0 = \hat{A} \hat{Q} + \hat{Q} \hat{A}^T + \hat{P}, \quad (3.1)$$

$$0 = \hat{A}^T \hat{P} + \hat{P} \hat{A} + \hat{R}, \quad (3.2)$$

$$0 = P_{12}^T Q_{12} + P_2 Q_2, \quad (3.3)$$

$$B_c = -(P_2^{-1} P_{12}^T Q_{12} + Q_{12}^T C^T V_2^{-1}), \quad (3.4)$$

$$C_c = -R_2^{-1} B^T (P_1 Q_{12} Q_2^{-1} + P_{12}), \quad (3.5)$$

where

$$\hat{P} \triangleq \begin{bmatrix} V_1 & 0 \\ 0 & B_c V_2 B_c^T \end{bmatrix}, \quad \hat{R} \triangleq \begin{bmatrix} R_1 & 0 \\ 0 & C_c^T R_2 C_c \end{bmatrix}$$

and $(n + n_c) \times (n + n_c)$ \hat{Q} , \hat{P} are partitioned into $n \times n$, $n \times n_c$, and $n_c \times n_c$ subblocks as

$$\hat{Q} = \begin{bmatrix} Q_1 & Q_{12} \\ Q_{12}^T & Q_2 \end{bmatrix}, \quad \hat{P} = \begin{bmatrix} P_1 & P_{12} \\ P_{12}^T & P_2 \end{bmatrix}.$$

Expanding (3.1) and (3.2) yields

$$0 = A Q_1 + Q_1 A^T + B C_c Q_{12}^T + Q_{12} (B C_c)^T + V_1, \quad (3.6)$$

$$0 = A Q_{12} + Q_{12} A^T + B C_c Q_2 + Q_1 (B C_c)^T, \quad (3.7)$$

$$0 = A_c Q_2 + Q_2 A_c^T + B_c C Q_{12} + Q_{12}^T (B_c C)^T + B_c V_2 B_c^T, \quad (3.8)$$

$$0 = A^T P_1 + P_1 A + (B C_c)^T P_{12}^T + P_{12} B_c C + R_1, \quad (3.9)$$

$$0 = P_{12} A_c + A^T P_{12} + (B C_c)^T P_2 + P_1 B C_c, \quad (3.10)$$

$$0 = A_c^T P_2 + P_2 A_c + (B C_c)^T P_{12} + P_{12}^T B C_c + C_c^T R_2 C_c. \quad (3.11)$$

Writing (3.8) as (see [26], [27])

$$0 = (A_c + B_c C Q_{12} Q_2^{-1}) Q_2 + Q_2 (A_c + B_c C Q_{12} Q_2^{-1})^T + B_c V_2 B_c^T$$

where Q_2^{-1} is the Moore–Penrose or Drazin generalized inverse of Q_2 , it follows from [28, Lemmas 2.1 and 12.2] that Q_2 is positive definite. Similarly, (3.11) implies that P_2 is positive definite. This justifies (3.4) and (3.5).

Now define the $n \times n$ nonnegative-definite matrices (see [26], [27])

$$Q \triangleq Q_1 - Q_{12} Q_2^{-1} Q_{12}^T, \quad P \triangleq P_1 - P_{12} P_2^{-1} P_{12}^T,$$

$$\hat{Q} \triangleq Q_{12} Q_2^{-1} Q_{12}^T, \quad \hat{P} \triangleq P_{12} P_2^{-1} P_{12}^T$$

and note that (3.3) implies (2.8) and (2.9) with

$$G \triangleq Q_2^{-1} Q_{12}^T, \quad M \triangleq Q_2 P_2, \quad \Gamma \triangleq -P_2^{-1} P_{12}^T.$$

Since $Q_2 P_2 = P_2^{-1/2} (P_2^{1/2} Q_2 P_2^{1/2}) P_2^{1/2}$, M is positive semisimple. Sylvester's inequality yields (2.17). Note also that

$$\hat{Q} = \tau \hat{Q}, \quad \hat{P} = \hat{P} \tau. \quad (3.12)$$

Next (2.11) and (2.12) follow from (3.4) and (3.5) by using the identities

$$Q_1 = Q + \hat{Q}, \quad P_1 = P + \hat{P}, \quad (3.13)$$

$$Q_{12} = \hat{Q} \Gamma^T, \quad P_{12} = -\hat{P} G^T, \quad (3.14)$$

$$Q_2 = \Gamma \hat{Q} \Gamma^T, \quad P_2 = G \hat{P} G^T. \quad (3.15)$$

Now substitute (2.11), (2.12), and (3.13)–(3.15) into (3.6)–(3.11) and use the relations

$$B_c C = \Gamma Q \Sigma, \quad B C_c = -\Sigma P G^T,$$

$$B_c V_2 B_c^T = \Gamma Q \Sigma Q \Gamma^T, \quad C_c^T R_2 C_c = G P \Sigma P G^T.$$

Then (2.10) follows from (3.8)–(3.7). Substituting (2.10) into (3.7), (3.8), (3.10), and (3.11) shows that $((3.7)G)^T$ and $-(3.10)\Gamma$ are precisely (2.15) and (2.16). Since $G^T(3.8)G = (2.15)\tau$ and $\Gamma^T(3.11)\Gamma = \tau(2.16)$, (3.8) and (3.11) can be omitted. Finally, using (3.12) it follows that (2.13) = (3.6) + (2.15) τ = (2.15)–(2.15) T and similarly for (2.14). \square

IV. DIRECTIONS FOR FURTHER RESEARCH

With regard to the existence of a stabilizing compensator, known results (e.g., [28]–[34]) can be exploited to a great extent. A numerical algorithm for solving the optimal projection equations has been developed in [24] and [35]. The proposed computational scheme is philosophically quite different from gradient search algorithms [2], [3], [6], [7], [9], [11], [36], [37] in that it operates through direct solution of the optimal projection equations by iterative refinement of the optimal projection. Methods for eliminating local extrema are being investigated by applying component cost analysis [17]. Generalizations of the optimal projection equations can arise by considering the following extensions of the fixed-order dynamic-compensation problem.

1) *Discrete-Time System/Discrete-Time Compensator:* Digital implementation can be modeled by a discrete-time compensator with control of a continuous-time system facilitated by sampling and reconstruction devices.

2) *Cross Weighting/Correlated Disturbance and Observation Noise*: This extension is straightforward and entirely analogous to the LQG case (see, e.g., [3, p. 351]).

3) *Singular Observation Noise/Singular Control Weighting*: With due attention to (2.7), direct output feedback can be used in the singular case. The nature of the problem forebodes all of the difficulties associated with the singular LQG problem. Note that the output feedback problem [22], [38], when viewed in this context, is *highly singular*.

4) *Infinite-Dimensional Systems*: The optimal projection equations have been extended in [39] and [40] to the case in which (2.1) is a distributed parameter system, for example, a partial or functional differential equation.

5) *Decentralized Fixed-Order Controller*: The optimal projection equations can be derived for the case in which the dynamic controller has a fixed decentralized structure.

6) *Parameter Uncertainties*: The original derivation in [23] treated a Stratonovich state-dependent noise model representing parameter uncertainties in the plant. Further consideration of control- and measurement-dependent noise raises the possibility of directly including the impact of parameter uncertainties in the design of robust, implementable compensation for large-order systems.

REFERENCES

- [1] T. L. Johnson and M. Athans, "On the design of optimal constrained dynamic compensators for linear constant systems," *IEEE Trans. Automat. Contr.*, vol. AC-15, pp. 658-660, 1970.
- [2] W. S. Levine, T. L. Johnson, and M. Athans, "Optimal limited state variable feedback controllers for linear systems," *IEEE Trans. Automat. Contr.*, vol. AC-16, pp. 785-793, 1971.
- [3] K. Kwakernaak and R. Sivan, *Linear Optimal Control Systems*. New York: Wiley, 1972.
- [4] D. B. Rom and P. E. Sarachik, "The design of optimal compensators for linear constant systems with inaccessible states," *IEEE Trans. Automat. Contr.*, vol. AC-18, pp. 509-512, 1973.
- [5] M. Sidiq and B.-Z. Kurtaran, "Optimal low-order controllers for linear stochastic systems," *Int. J. Contr.*, vol. 22, pp. 377-387, 1975.
- [6] J. M. Mendel and J. Feather, "On the design of optimal time-invariant compensators for linear stochastic time-invariant systems," *IEEE Trans. Automat. Contr.*, vol. AC-20, pp. 653-657, 1975.
- [7] S. Basuthakur and C. H. Knapp, "Optimal constant controllers for stochastic linear systems," *IEEE Trans. Automat. Contr.*, vol. AC-20, pp. 664-666, 1975.
- [8] R. B. Asher and J. C. Durrett, "Linear discrete stochastic control with a reduced-order dynamic compensator," *IEEE Trans. Automat. Contr.*, vol. AC-21, pp. 626-627, 1976.
- [9] W. J. Naeije and O. H. Bosgra, "The design of dynamic compensators for linear multivariable systems," in *Proc. 1977 IFAC*, Fredrickton, N. B., Canada, 1977, pp. 205-212.
- [10] P. J. Blavillain and T. L. Johnson, "Invariants of optimal minimal-order observer based compensators," *IEEE Trans. Automat. Contr.*, vol. AC-23, pp. 473-474, 1978.
- [11] C. J. Wenk and C. H. Knapp, "Parameter optimization in linear systems with arbitrarily constrained controller structure," *IEEE Trans. Automat. Contr.*, vol. AC-25, pp. 496-500, 1980.
- [12] J. O'Reilly, "Optimal low-order feedback controllers for linear discrete-time systems," in *Control and Dynamic Systems*, Vol. 16, C. T. Leondes, Ed. New York: Academic, 1980.
- [13] B. D. Moore, "Principal component analysis in linear systems: Controllability, observability and model reduction," *IEEE Trans. Automat. Contr.*, vol. AC-26, pp. 17-32, 1981.
- [14] E. I. Verriest, "Suboptimal LQG-design and balanced realizations," in *Proc. 20th IEEE Conf. Decision Contr.*, San Diego, CA, Dec. 1981, pp. 686-687.
- [15] L. Pernebo and L. M. Silverman, "Model reduction via balanced state space representations," *IEEE Trans. Automat. Contr.*, vol. AC-27, pp. 382-387, 1982.
- [16] E. A. Jonckheere and L. M. Silverman, "A new set of invariants for linear systems-application to reduced-order compensator design," *IEEE Trans. Automat. Contr.*, vol. AC-28, pp. 953-964, 1983.
- [17] A. Yousuff and R. E. Skelton, "Controller reduction by component cost analysis," *IEEE Trans. Automat. Contr.*, vol. AC-29, pp. 520-530, 1984.
- [18] D. C. Hyland and D. S. Bernstein, "The optimal projection approach to model reduction and the relationship between the methods of Wilson and Moore," in *23rd IEEE Conf. Decision Contr.*, Las Vegas, NV, Dec. 1984.
- [19] D. A. Wilson, "Optimum solution of model-reduction problem," *Proc. IEE*, vol. 117, pp. 1161-1165, 1970.
- [20] C. R. Rao and S. K. Mitra, *Generalized Inverse of Matrices and its Applications*. New York: Wiley, 1971.
- [21] S. L. Campbell and C. D. Meyer, Jr., *Generalized Inverses of Linear Transformations*. London: Pitman, 1979.
- [22] J. Medanic, "On stabilization and optimization by output feedback," in *Proc. 12th Annual Asilomar Conf. Circuits and Syst.*, 1978, pp. 412-416.
- [23] D. C. Hyland, "Optimality conditions for fixed-order dynamic compensation of flexible spacecraft with uncertain parameters," in *Proc. AIAA 20th Aerospace Sciences Meet.*, Orlando, FL, Jan. 1982, paper 82-0312.
- [24] D. C. Hyland, "The optimal projection approach to fixed-order compensation: Numerical methods and illustrative result," in *AIAA 21st Aerospace Sciences Meet.*, Reno, NV, Jan. 1983, paper 83-0303.
- [25] D. C. Hyland and D. S. Bernstein, "Explicit optimality conditions for fixed-order dynamic compensation," in *Proc. 22nd IEEE Conf. Decision Contr.*, San Antonio, TX, Dec. 1983.
- [26] A. Albert, "Conditions for positive and nonnegative definiteness in terms of pseudo inverse," *SIAM J. Appl. Math.*, vol. 17, pp. 434-440, 1969.
- [27] E. Kreindler and A. Jameson, "Conditions for nonnegativeness of partitioned matrices," *IEEE Trans. Automat. Contr.*, vol. AC-17, pp. 147-148, 1972.
- [28] W. M. Wonham, *Linear Multivariable Control. A Geometric Approach*. New York: Springer-Verlag, 1974.
- [29] F. M. Brash and J. B. Pearson, "Pole placement using dynamic compensators," *IEEE Trans. Automat. Contr.*, vol. AC-15, pp. 34-43, 1970.
- [30] R. Ahmari and A. G. Vachoux, "On the pole assignment in linear systems with fixed-order compensators," *Int. J. Contr.*, vol. 17, pp. 397-404, 1973.
- [31] D. C. Youla, J. J. Bongiorno, Jr., and C. N. Lu, "Single-loop feedback-stabilization of linear multivariable dynamical plants," *Automatica*, vol. 10, pp. 159-173, 1974.
- [32] H. Seraji, "An approach to dynamic compensator design for pole assignment," *Int. J. Contr.*, vol. 21, pp. 955-966, 1975.
- [33] R. V. Patel, "Design of dynamic compensators for pole assignment," *Int. J. Syst. Sci.*, vol. 7, pp. 207-224, 1976.
- [34] C. A. Desoer, R. W. Liu, J. Murrar, and R. Sacks, "Feedback system design: The fractional representation approach to analysis and synthesis," *IEEE Trans. Automat. Contr.*, vol. AC-25, pp. 399-412, 1980.
- [35] D. C. Hyland, "Comparison of various controller-reduction methods: Suboptimal versus optimal projection," in *Proc. AIAA Dynamics Specialists Conf.*, Palm Springs, CA, May 1984, pp. 381-389.
- [36] H. R. Sirisena and S. S. Choi, "Design of optimal constrained dynamic compensators for non-stationary linear stochastic systems," *Int. J. Contr.*, vol. 25, pp. 513-524, 1977.
- [37] D. P. Looze and N. R. Sandell, Jr., "Gradient calculations for linear quadratic fixed control structure problems," *IEEE Trans. Automat. Contr.*, vol. AC-25, pp. 285-288, 1980.
- [38] W. S. Levine and M. Athans, "On the determination of the optimal constant output feedback gains for linear multivariable systems," *IEEE Trans. Automat. Contr.*, vol. AC-15, pp. 44-48, 1970.
- [39] D. S. Bernstein and D. C. Hyland, "The optimal projection equations for fixed-order dynamic compensation of distributed parameter systems," in *Proc. AIAA Dynamics Specialists Conf.*, Palm Springs, CA, May 1984, pp. 396-400.
- [40] D. S. Bernstein, "Explicit optimality conditions for finite-dimensional fixed-order dynamic compensation of infinite-dimensional systems," submitted for publication.

APPENDIX C: Distributed Parameter Systems

- 37. D. S. Bernstein and D. C. Hyland, "The Optimal Projection Equations for Finite-Dimensional Fixed-Order Dynamic Compensation of Infinite-Dimensional Systems," SIAM J. Contr. Optim., Vol. 24, pp. 122-151, 1986.
- 122. D. S. Bernstein and I. G. Rosen, "Finite-Dimensional Approximation for Optimal Fixed-Order Compensation of Distributed Parameter Systems," submitted to Opt. Contr. Appl. Meth.

THE OPTIMAL PROJECTION EQUATIONS FOR FINITE-DIMENSIONAL FIXED-ORDER DYNAMIC COMPENSATION OF INFINITE-DIMENSIONAL SYSTEMS*

DENNIS S. BERNSTEIN† AND DAVID C. HYLAND†

Abstract. One of the major difficulties in designing implementable finite-dimensional controllers for distributed parameter systems is that such systems are inherently infinite dimensional while controller dimension is severely constrained by on-line computing capability. While some approaches to this problem initially seek a correspondingly infinite-dimensional control law whose finite-dimensional approximation may be of impractically high order, the usual engineering approach involves first approximating the distributed parameter system with a high-order discretized model followed by design of a relatively low-order dynamic controller. Among the numerous approaches suggested for the latter step are model/controller reduction techniques used in conjunction with the standard *LQG* result. An alternative approach, developed in [36], relies upon the discovery in [31] that the necessary conditions for optimal fixed-order dynamic compensation can be transformed into a set of equations possessing remarkable structural coherence. The present paper generalizes this result to apply directly to the distributed parameter system itself. In contrast to the pair of operator Riccati equations for the "full-order" *LQG* case, the optimal finite-dimensional fixed-order dynamic compensator is characterized by *four* operator equations (two modified Riccati equations and two modified Lyapunov equations) coupled by an oblique projection whose rank is precisely equal to the order of the compensator and which determines the optimal compensator gains. This "optimal projection" is obtained by a full-rank factorization of the product of the finite-rank nonnegative-definite Hilbert-space operators which satisfy the pair of modified Lyapunov equations. The coupling represents a graphic portrayal of the demise of the classical separation principle for the finite-dimensional reduced-order controller case. The results obtained apply to a semigroup formulation in Hilbert space and thus are applicable to control problems involving a broad range of specific partial and functional differential equations.

Key words. optimality conditions, finite-dimensional fixed-order dynamic compensator, infinite-dimensional system, distributed parameter system, semisimple operator, oblique projection, Drazin generalized inverse

1. Introduction. One of the major difficulties in designing active controllers for distributed parameter systems is that such systems are inherently infinite dimensional while implementable controllers are necessarily finite dimensional with controller dimension severely constrained by on-line computing capability. As pointed out by Balas ([1], see also [2]), control design for distributed parameter systems entails the *practical constraints* of 1) finitely many sensors and actuators, 2) a finite-dimensional controller and 3) natural system dissipation. The validity of 2) is apparent from the fact that processing and transmitting electrical signals by conventional analog or digital components constitutes finite-dimensional action. Although distributed parameter devices can also be utilized, their fabrication and implementation can incorporate at most a finite number of design specifications.¹ Hence, although distributed parameter systems are most accurately represented by infinite-dimensional models, real-world

* Received by the editors December 6, 1983, and in revised form September 15, 1984. This work was performed at Lincoln Laboratory/MIT and was sponsored by the Department of the Air Force.

† Harris Corporation, Government Aerospace Systems Division, Controls Analysis and Synthesis Group, Melbourne, Florida 32901.

¹ Examples of such components include tapped delay lines and surface acoustic wave devices. Although acoustoelectric convolvers [3, p. 465] can perform continuous-time integration, synthesis of the desired impulse-response kernel can incorporate only finitely many specified parameters. The obvious fact should also be noted that physical limitations impose an upper bound on the number of design parameters that can be incorporated in the construction of any device. For an extensive treatment of this subject, see [72].

constraints require that implementable controllers be modelled as lumped parameter systems.

Clearly, the above observations effectively preclude the possibility of realizing infinite-dimensional controllers that involve full-state feedback or full-state estimation (see, e.g., [4]–[6] and the numerous references therein). Although finite-dimensional approximation schemes have been applied to optimal infinite-dimensional control laws ([7]–[9]), these results only guarantee optimality in the limit, i.e., as the order of the approximating controller increases without bound. Hence, there is no guarantee that a particular approximate (i.e., discretized) controller is actually optimal over the class of approximate controllers of a given order dictated by implementation constraints. Moreover, even if an optimal *approximate* finite-dimensional controller could be obtained, it would almost certainly be suboptimal in the class of *all* controllers of the given order.

Although the usual engineering approach to this problem is to replace the distributed parameter system with a high-order finite-dimensional model, analogous, fundamental difficulties remain since application of *LQG* leads to a controller whose order is identical to that of the high-order approximate model. Attempts to remedy this problem usually rely upon some method of open-loop model reduction or closed-loop controller reduction (see, e.g., [10]–[15]). Most of these techniques (with the exception of [11]) are ad hoc in nature, however, and hence guarantees of optimality and stability may be lacking.

A more direct approach that avoids both model and controller reduction is to fix the controller structure and optimize the performance criterion with respect to the controller parameters. Although much effort was devoted to this approach (see, e.g., [16]–[30]), progress in this direction was impeded by the extreme complexity of the nonlinear matrix equations arising from the first-order necessary conditions. What was lacking, to quote the insightful remarks of [24], was a “deeper understanding of the structural coherence of these equations.” The key to unlocking these unwieldy equations was subsequently discovered by Hyland in [31] and developed in [32]–[36]. Specifically, it was found that these equations harbored the definition of an *oblique projection* (i.e., idempotent matrix) which is a consequence of optimality and *not* the result of an ad hoc assumption. By exploiting the presence of this “optimal projection,” the originally *very* complex stationary conditions can be transformed without loss of generality into much simpler and more tractable forms. The resulting equations (see [36, (2.10)–(2.17)]) preserve the simple form of *LQG* relations for the gains in terms of covariance and cost matrices which, in turn, are determined by a coupled system of two modified Riccati equations and two modified Lyapunov equations. This coupling, by means of the optimal projection, represents a graphic portrayal of the demise of the classical separation principle for the reduced-order controller case. When, as a special case, the order of the compensator is required to be equal to the order of the plant, the modified Riccati equations immediately reduce to the standard *LQG* Riccati equations and the modified Lyapunov equations express the proviso that the compensator be minimal, i.e., controllable and observable. Since the *LQG* Riccati equations as such are nothing more than the necessary conditions for full-order compensation, the “optimal projection equations” appear to provide a clear and simple generalization of standard *LQG* theory.

The fact that the optimal projection equations consist of *four* coupled matrix equations, i.e., two modified Riccati equations and two modified Lyapunov equations, can readily be explained by the following simple reason. Reduced-order control-design methods often involve either *LQG* applied to a reduced-order model or model reduction

applied to a full-order *LQG* design, and hence both approaches require the solution of precisely four equations: two Riccati equations (for *LQG*) plus two Lyapunov equations (for system reduction via balancing, as in [12], [14]). The *coupled* form of the optimal projection equations is thus a strong reminder that the *LQG* and order-reduction operations *cannot* be iterated but must, in a precise sense, be performed *simultaneously*. This situation is partly due to the fact that the optimal projection matrix may not be of the form $\begin{bmatrix} I & 0 \\ 0 & 0 \end{bmatrix}$ even in the basis corresponding to the "balanced" realization [12], [14]. This point is explored in [37], [37a] where the solution to the optimal model-reduction problem is characterized by a pair of modified Lyapunov equations which are also coupled by an oblique projection.

Returning now to the distributed parameter problem, it should be mentioned that notable exceptions to the previously mentioned work on distributed parameter controllers are the contributions of Johnson [38] and Pearson [39], [40] who suggest fixing the order of the finite-dimensional compensator while retaining the distributed parameter model. Progress in this direction, however, was impeded not only by the intractability of the optimality conditions that were available for the finite-dimensional problem (as in [16]–[30]), but also by the lack of a suitable generalization of these conditions to the infinite-dimensional case. The purpose of the present paper is to make significant progress in filling these gaps, i.e., by deriving explicit optimality conditions which directly characterize the optimal finite-dimensional fixed-order dynamic compensator for an infinite-dimensional system and which are exactly analogous to the highly simplified optimal projection equations obtained in [31]–[34], [36] for the finite-dimensional case. Specifically, instead of a system for four matrix equations we obtain a system of four *operator* equations whose solutions characterize the optimal finite-dimensional fixed-order dynamic compensator. Moreover, the optimal projection now becomes a bounded idempotent Hilbert-space operator whose rank is precisely equal to the order of the compensator.

The mathematical setting we use is standard: a linear time-invariant differential system in Hilbert space with additive white noise, finitely many controls and finitely many noisy measurements (thus satisfying the first practical constraint mentioned above). The input and output maps are assumed to be bounded. Since the only explicit assumption on the unbounded dynamics operator is that it generate a strongly continuous semigroup, the results are potentially applicable to a broad range of specific partial and functional differential equations. The *actual* applicability of our results is essentially limited by practical constraint 3). Since we are concerned with the steady-state problem, we implicitly assume that the distributed parameter system is stabilizable, i.e., that there exists a dynamic compensator of a given order such that the closed-loop system is uniformly stable. We note that stabilizing compensators do exist for the wide class of problems considered in [41] and [42] which includes delay, parabolic and damped hyperbolic systems. The question of *how much* damping is required for stabilizability of hyperbolic systems is a crucial issue in designing controllers for large flexible space structures [7], [43]–[49a].

It is important to point out that the results of this paper can immediately be specialized to finite-dimensional systems by requiring that the Hilbert space characterizing the dynamical system be finite-dimensional. Then all unboundedness considerations can be ignored, adjoints can be interpreted as transposes and other obvious simplifications can be invoked. The only mathematical aspect requiring attention is the treatment of white noise which, for general handling of the infinite-dimensional case, is interpreted according to [6].² For the finite-dimensional case, however, the standard

classical notions suffice and the results go through with virtually no modifications.

The contents of the paper are as follows. Section 2 contains preliminary notation in addition to particular results for use later in the paper. Section 3 presents the optimal steady-state finite-dimensional fixed-order dynamic-compensation problem and the Main Theorem gives the necessary conditions in the form of the optimal projection equations (3.15)–(3.18). We then develop a series of results which serve to elucidate several aspects of the Main Theorem. Section 4 is devoted to the proof of the Main Theorem. The reader is alerted to the two crucial steps required. The first step involves generalizing to the infinite-dimensional case the derivation of the necessary conditions in their "primitive" form (see (4.27)–(4.29) and (4.48)–(4.53)). The derivation in [31]–[33], [36] involving Lagrange multipliers is invalid in the infinite-dimensional case due to the presence of the unbounded system-dynamics operator. Instead, we use the gramian form of the closed-loop covariance operator to obtain a dual problem formulation and then proceed to derive the primitive necessary conditions by means of a lengthy, but direct, computation (Lemma 4.7). The second crucial step involves transforming the primitive form of the necessary conditions to the final form given in the Main Theorem. This laborious computation was first carried out in [31], [32] and was subsequently facilitated in [33], [36] by means of a judicious change of variables (see (4.32), (4.33)). Finally, some concluding remarks are given in § 5.

2. Preliminaries. In this section we introduce general notation along with basic definitions and results for use in later sections. Our principal references are [6], [50] and [51].

Throughout this section let \mathcal{H} , \mathcal{H}' and \mathcal{H}'' denote real separable Hilbert spaces with norm $\|\cdot\|$ and inner product (\cdot, \cdot) and let $\mathcal{B}(\mathcal{H}, \mathcal{H}')$ denote the space of bounded linear operators from \mathcal{H} into \mathcal{H}' . For $L \in \mathcal{B}(\mathcal{H}, \mathcal{H}')$, $\|L\|$ is the norm of L , $\mathcal{R}(L)$ is the range of L , $\mathcal{N}(L)$ is the null space of L , $\rho(L)$ is the rank of L (set $\rho(L) = \infty$ if L does not have finite rank), L^{-1} is the inverse of L when L is invertible, i.e., when L has a bounded inverse, L^* is the adjoint of L and $L^{-*} \triangleq (L^*)^{-1}$. Recall that $\|L\| = \|L^*\|$ and that $\rho(L) = \rho(L^*)$ [50, p. 161]. Now suppose that $\mathcal{H} = \mathcal{H}'$ so that $L \in \mathcal{B}(\mathcal{H}) \triangleq \mathcal{B}(\mathcal{H}, \mathcal{H})$. If $LL^* = L^*L$ then L is *normal* and if $L = L^*$ then L is *selfadjoint*. If L is selfadjoint and $(Lx, x) \geq 0$, $x \in \mathcal{H}$, then L is *nonnegative definite*. Note that the selfadjointness assumption is included in the definition since the Hilbert spaces are assumed real. If L is nonnegative definite then $L^{1/2}$ denotes the (unique) nonnegative-definite square root of L . Call L *semisimple* (resp., *real semisimple*, *nonnegative semisimple*) if there exists invertible $S \in \mathcal{B}(\mathcal{H})$ such that SLS^{-1} is normal (resp., selfadjoint, nonnegative definite). This implies that SLS^{-1} has a complete set of orthonormal eigenvectors and, in the real-semisimple or nonnegative-semisimple cases, has real or nonnegative eigenvalues.

Recall that if $L \in \mathcal{B}(\mathcal{H})$ is compact then L has at most a countable number of eigenvalues and all nonzero eigenvalues have finite multiplicity. Hence, for $L \in \mathcal{B}(\mathcal{H}, \mathcal{H}')$ compact, let $\{\alpha_i\}$ be the (at most countable) sequence of eigenvalues of $(LL^*)^{1/2}$ with appropriate multiplicity and $\alpha_1 \geq \alpha_2 \geq \dots > 0$ [50, p. 261]. Then $\mathcal{B}_1(\mathcal{H}, \mathcal{H}')$ denotes the set of *trace class* (or *nuclear*) operators, i.e., the set of compact

² Alternatively, we could have adopted the white noise formulation of [4]. The main difference between the two white noise formalisms is that Balakrishnan works with finitely additive rather than countably additive measures. Strictly speaking, then, even in finite dimensions Balakrishnan's white noise is different from the standard notion (see [6, pp. 307, 315]).

$L \in \mathcal{B}(\mathcal{H}, \mathcal{H}')$ for which $\sum_i \alpha_i < \infty$ [50, p. 521]. $\mathcal{B}_1(\mathcal{H}, \mathcal{H}')$ is a Banach space with norm

$$\|L\|_1 \triangleq \sum_i \alpha_i.$$

If $\sum_i \alpha_i^2 < \infty$ then $L \in \mathcal{B}_2(\mathcal{H}, \mathcal{H}')$, the set of Hilbert-Schmidt operators, which is a Banach space with norm

$$\|L\|_2 \triangleq \left[\sum_i \alpha_i^2 \right]^{1/2}.$$

Note that $\|L\| \leq \|L\|_2 \leq \|L\|_1$, $\|L\| = \|L^*\|$, $\|L\|_1 = \|L^*\|_1$ and $\|L\|_2 = \|L^*\|_2$. If $\mathcal{H} = \mathcal{H}'$, then we write $\mathcal{B}_1(\mathcal{H})$ and $\mathcal{B}_2(\mathcal{H})$ for $\mathcal{B}_1(\mathcal{H}, \mathcal{H})$ and $\mathcal{B}_2(\mathcal{H}, \mathcal{H})$, respectively. Note that if nonnegative-definite $L \in \mathcal{B}_1(\mathcal{H})$ then $L^{1/2} \in \mathcal{B}_2(\mathcal{H})$.

If $L \in \mathcal{B}_1(\mathcal{H}, \mathcal{H}')$ and $S \in \mathcal{B}(\mathcal{H}', \mathcal{H}'')$ then

$$\|SL\|_1 \leq \|S\| \|L\|_1$$

and hence $SL \in \mathcal{B}_1(\mathcal{H}, \mathcal{H}'')$. Similarly, under suitable hypotheses,

$$\|LS\|_1 \leq \|S\| \|L\|_1,$$

and

$$\|SL\|_1 \leq \|S\|_2 \|L\|_2.$$

LEMMA 2.1. Suppose $L \in \mathcal{B}_1(\mathcal{H})$ and let $\{\lambda_i\}$ denote the nonzero eigenvalues of L with appropriate multiplicity. Then [51, p. 89]

$$\sum_i |\lambda_i| \leq \|L\|_1.$$

If L is selfadjoint then [50, p. 522]

$$\sum_i |\lambda_i| = \|L\|_1.$$

If L is nonnegative definite then

$$\sum_i \lambda_i = \|L\|_1.$$

Let $L \in \mathcal{B}_1(\mathcal{H})$. Then define [50, p. 523] the trace functional $\text{tr}: \mathcal{B}_1(\mathcal{H}) \rightarrow \mathbb{R}$ by

$$\text{tr } L \triangleq \sum_i \langle L\phi_i, \phi_i \rangle,$$

where the summation is independent of the choice of orthonormal basis $\{\phi_i\}$. The trace satisfies $\text{tr } L = \text{tr } L^*$, $\text{tr } SL = \text{tr } LS$ for all $S \in \mathcal{B}(\mathcal{H})$, $\text{tr } ST = \text{tr } TS$ for all $S, T \in \mathcal{B}_2(\mathcal{H})$ and $\text{tr } (\alpha T + \beta S) = \alpha(\text{tr } T) + \beta(\text{tr } S)$ for all $\alpha, \beta \in \mathbb{R}$ and $S, T \in \mathcal{B}_1(\mathcal{H})$.

LEMMA 2.2. Suppose $L \in \mathcal{B}_1(\mathcal{H})$ and let $\{\lambda_i\}$ denote the nonzero eigenvalues of L with appropriate multiplicity. Then [51, p. 139]

$$\text{tr } L = \sum_i \lambda_i$$

and hence (by Lemma 2.1)

$$|\text{tr } L| \leq \|L\|_1.$$

If L is nonnegative definite then

$$\text{tr } L = \|L\|_1.$$

COROLLARY 2.1. For each $S \in \mathcal{B}(\mathcal{H})$ the linear functionals

$$L \rightarrow \text{tr } SL: \mathcal{B}_1(\mathcal{H}) \rightarrow \mathbb{R},$$

$$L \rightarrow \text{tr } LS: \mathcal{B}_1(\mathcal{H}) \rightarrow \mathbb{R}$$

are continuous. For each $L \in \mathcal{B}_1(\mathcal{H})$ the linear functionals

$$S \rightarrow \text{tr } LS: \mathcal{B}(\mathcal{H}) \rightarrow \mathbb{R},$$

$$S \rightarrow \text{tr } SL: \mathcal{B}(\mathcal{H}) \rightarrow \mathbb{R}$$

are continuous.

Although showing that a bounded linear operator is trace class is slightly more involved than the above characterizations of $\mathcal{B}_1(\mathcal{H})$, the following result will suffice for our purposes (see [52, p. 96], or [52a, p. 171]).

LEMMA 2.3. Let $L \in \mathcal{B}(\mathcal{H})$ be nonnegative definite. Then

$$\sum_i \langle L\phi_i, \phi_i \rangle,$$

whether finite or infinite, is independent of the orthonormal basis $\{\phi_i\}$. The summation is finite if and only if $L \in \mathcal{B}_1(\mathcal{H})$.

Many of the operators introduced in the following section have finite-dimensional domain or range space and hence are degenerate, i.e., have finite rank. Recall that degenerate operators are necessarily trace class. The following result, which generalizes [53, Thm. 2.1, p. 240] in certain respects, will be fundamental in decomposing finite-rank operators.

LEMMA 2.4. Suppose $L_1, \dots, L_r \in \mathcal{B}(\mathcal{H}, \mathcal{H}')$ have finite rank. Then there exists a finite-dimensional subspace $\mathcal{M} \subset \mathcal{H}$ such that $L_i \mathcal{M}^\perp = 0$, $i = 1, \dots, r$. Furthermore, if $\mathcal{H} = \mathcal{H}'$ then \mathcal{M} can be chosen such that $L_i \mathcal{M} \subset \mathcal{M}$, $i = 1, \dots, r$.

Proof. It suffices to consider the case $r = 1$. Writing L for L_1 , note that since $\rho(L^*) < \infty$, $\mathcal{N}(L)^\perp = \mathcal{R}(L^*)$ [50, p. 155] and $\mathcal{N}(L)$ is closed, the first statement holds with $\mathcal{M} = \mathcal{N}(L)^\perp$. When $\mathcal{H} = \mathcal{H}'$ set $\mathcal{M} = \mathcal{N}(L)^\perp + \mathcal{R}(L)$ and note that $\mathcal{M}^\perp = \mathcal{N}(L) \cap \mathcal{R}(L)^\perp \subset \mathcal{N}(L)$ and $L\mathcal{M} \subset \mathcal{R}(L) \subset \mathcal{M}$. \square

The following generalization of Sylvester's inequality [54, p. 66] will be used repeatedly in handling finite-rank operators.

LEMMA 2.5. Let $L \in \mathcal{B}(\mathcal{H}, \mathcal{H}')$ and $S \in \mathcal{B}(\mathcal{H}', \mathcal{H}'')$. Then

$$(2.1) \quad \rho(SL) \leq \min \{\rho(S), \rho(L)\}.$$

If $\dim \mathcal{H}' = \nu < \infty$, then

$$(2.2) \quad \rho(S) + \rho(L) - \nu \leq \rho(SL).$$

Proof. If either S or L does not have finite rank then (2.1) is immediate. If both S and L have finite rank then the standard arguments [54] used to prove the finite-dimensional version of (2.1) remain valid. To prove (2.2), note that Lemma 2.4 implies that there exist orthonormal bases for \mathcal{H} and \mathcal{H}' with respect to which L has the matrix representation $\begin{bmatrix} \tilde{L} & 0 \\ 0 & 0 \end{bmatrix}$, where $\tilde{L} \in \mathbb{R}^{\nu \times p}$. Similarly, there exist orthonormal bases for \mathcal{H}' and \mathcal{H}'' with respect to which S has the matrix representation $\begin{bmatrix} \tilde{S} & 0 \\ 0 & 0 \end{bmatrix}$, where $\tilde{S} \in \mathbb{R}^{q \times \nu}$. Since the two cited bases for \mathcal{H}' may be different, let orthogonal $U \in \mathbb{R}^{\nu \times \nu}$ be the matrix representation (with respect to either basis for \mathcal{H}') for the change in orthonormal basis [6, p. 100]. Hence SL has the matrix representation $\begin{bmatrix} \tilde{S}U\tilde{L} & 0 \\ 0 & 0 \end{bmatrix}$ and (2.2) follows from the known result [54, p. 66]. \square

As in the proof of Lemma 2.5, we shall utilize the infinite-matrix representation of an operator with respect to an orthonormal basis. All matrix representations given

here will consist of real entries since the Hilbert spaces involved are real. When the orthonormal bases are specified and no confusion can arise, we shall not differentiate between an operator and its matrix representation. We shall use the infinite identity matrix I_∞ interchangeably with the identity $I_{\mathcal{H}}$ on \mathcal{H} .

When dealing with finite-dimensional Euclidean spaces the notation and terminology introduced above will be utilized with only minor changes. For example, bounded linear operators will be represented by matrices whose elements are determined according to fixed orthonormal bases and hence we identify $\mathcal{B}(\mathcal{H}, \mathcal{H}) = \mathcal{B}(\mathbb{R}^n, \mathbb{R}^n)$. Note that if $L \in \mathcal{B}(\mathbb{R}^n, \mathbb{R}^m)$ and $S \in \mathcal{B}(\mathbb{R}^m, \mathbb{R}^n)$ then SL is an $m \times n$ matrix which is independent of any particular orthonormal basis for \mathcal{H} . The transposes of $x \in \mathbb{R}^n \triangleq \mathbb{R}^{n \times 1}$ and $M \in \mathbb{R}^{m \times n}$ are denoted by x^T and M^T and $M^{-T} \triangleq (M^T)^{-1}$. Let I_n denote the $n \times n$ identity matrix.

To specialize some of the above operator terminology to matrices, let $M \in \mathbb{R}^{n \times n}$. We shall say M is *nonnegative* (resp., *positive*) *diagonal* if M is diagonal with nonnegative (resp., positive) diagonal elements. M is *nonnegative* (resp., *positive*) *definite* if M is symmetric and $x^T M x \geq 0$ (resp., $x^T M x > 0$), $x \in \mathbb{R}^n$. Recall that M is symmetric (resp., nonnegative definite, positive definite) if and only if there exists orthogonal $U \in \mathbb{R}^{n \times n}$ such that UMU^T is diagonal (resp., nonnegative diagonal, positive diagonal). M is *semisimple* [55, p. 13], or *nondefective* [56, p. 375], if M has n linearly independent eigenvectors, i.e., M has a diagonal Jordan canonical form over the complex field. M is *real* (resp., *nonnegative*, *positive*) *semisimple* if M is semisimple with real (resp., nonnegative, positive) eigenvalues. Note that M is real (resp., nonnegative, positive) semisimple if and only if there exists invertible $S \in \mathbb{R}^{n \times n}$ such that SMS^{-1} is diagonal (resp., nonnegative diagonal, positive diagonal). Alternatively, M is real (resp., nonnegative, positive) semisimple if and only if there exists invertible $S \in \mathbb{R}^{n \times n}$ such that SMS^{-1} is symmetric (resp., nonnegative definite, positive definite).

LEMMA 2.6. *The product of two nonnegative- (resp., positive-) definite matrices is nonnegative (resp., positive) semisimple.*

Proof. If $S, L \in \mathbb{R}^{n \times n}$ are both nonnegative (resp., positive) definite then by [55, Thm. 6.2.5, p. 123] there exists invertible $\phi \in \mathbb{R}^{n \times n}$ such that $D_S \triangleq \phi^{-1} S \phi^{-T}$ and $D_L \triangleq \phi^T L \phi$ are nonnegative (resp., positive) diagonal. Hence, $SL = \phi D_S D_L \phi^{-1}$ is nonnegative (resp., positive) semisimple, as desired. Alternatively, if either S or L is positive definite, then the result follows from $SL = L^{-1/2} (L^{1/2} S L^{1/2}) L^{1/2}$ if L is positive definite or $SL = S^{1/2} (S^{1/2} L S^{1/2}) S^{-1/2}$ if S is positive definite. \square

3. Problem statement and the Main Theorem. We consider the following steady-state fixed-order dynamic-compensation problem. Given the dynamical system on $[0, \infty)$

$$(3.1) \quad \dot{x}(t) = Ax(t) + Bu(t) + H_1 w(t),$$

$$(3.2) \quad y(t) = Cx(t) + H_2 w(t),$$

design a finite-dimensional fixed-order dynamic compensator

$$(3.3) \quad \dot{x}_c(t) = A_c x_c(t) + B_c y(t),$$

$$(3.4) \quad u(t) = C_c x_c(t)$$

which minimizes the steady-state performance criterion

$$(3.5) \quad J(A_\infty, B_\infty, C_\infty) \triangleq \lim_{t \rightarrow \infty} E[(R_1 x(t), x(t)) + u(t)^T R_2 u(t)].$$

The following data are assumed. The state $x(t)$ is an element of a real separable Hilbert space \mathcal{X} and the state differential equation is interpreted in the weak sense (see, e.g., [6, pp. 229, 317]). The closed, densely defined operator $A: \mathcal{D}(A) \subset \mathcal{X} \rightarrow \mathcal{X}$ generates a strongly continuous semigroup e^{At} , $t \geq 0$. The control $u(t) \in \mathbb{R}^m$, $B \in \mathcal{B}(\mathbb{R}^m, \mathcal{X})$ and the operator $R_1 \in \mathcal{B}_1(\mathcal{X})$ and the matrix $R_2 \in \mathbb{R}^{m \times m}$ are nonnegative definite and positive definite, respectively. $w(\cdot)$ is a zero-mean Gaussian "standard white noise process" in $L_2((0, \infty), \mathcal{X})$ (see [6, p. 314]), where \mathcal{X}' is a real separable Hilbert space, $H_1 \in \mathcal{B}_2(\mathcal{X}', \mathcal{X})$, $H_2 \in \mathcal{B}(\mathcal{X}', \mathbb{R}^1)$ and "E" denotes expectation. We assume that $H_1 H_2^* = 0$, i.e., the disturbance and measurement noises are independent,³ and that $V_2 \triangleq H_2 H_2^* \in \mathbb{R}^1$ is positive definite, i.e., all measurements are noisy. Note that $V_1 \triangleq H_1 H_1^* \in \mathcal{B}_1(\mathcal{X})$ is nonnegative definite and trace class.⁴ The initial state $x(0)$ is Gaussian and independent of $w(\cdot)$. The observation $y(t) \in \mathbb{R}^1$ and $C \in \mathcal{B}(\mathcal{X}, \mathbb{R}^1)$. The dimension of the compensator state $x_c(t)$ is of fixed, finite order $n_c \leq \dim \mathcal{X}$ and the optimization is performed over $A_c \in \mathbb{R}^{n_c \times n_c}$, $B_c \in \mathbb{R}^{n_c \times 1}$ and $C_c \in \mathbb{R}^{1 \times n_c}$.

To handle the closed-loop system (3.1)–(3.4), we introduce the augmented state space $\tilde{\mathcal{X}} \triangleq \mathcal{X} \oplus \mathbb{R}^{n_c}$ which is a real separable Hilbert space with inner product $\langle \tilde{x}_1, \tilde{x}_2 \rangle \triangleq \langle x_1, x_2 \rangle + x_{c1}^T x_{c2}$, $\tilde{x}_i \triangleq (x_i, x_{ci})$. An operator $L \in \mathcal{B}(\tilde{\mathcal{X}})$ has a "decomposition" into operators $L_1 \in \mathcal{B}(\mathcal{X})$, $L_{12} \in \mathcal{B}(\mathbb{R}^{n_c}, \mathcal{X})$, $L_{21} \in \mathcal{B}(\mathcal{X}, \mathbb{R}^{n_c})$ and $L_2 \in \mathbb{R}^{n_c \times n_c}$ in the sense that for $\tilde{x} \triangleq (x, x_c) \in \tilde{\mathcal{X}}$, $L\tilde{x} = (L_1 x + L_{12} x_c, L_{21} x + L_2 x_c)$, or, in "block" form,

$$L = \begin{bmatrix} L_1 & L_{12} \\ L_{21} & L_2 \end{bmatrix}.$$

For later use note that

$$\|L\| \leq \|L_1\| + \|L_{12}\| + \|L_{21}\| + \|L_2\|$$

and

$$L^* = \begin{bmatrix} L_1^* & L_{21}^* \\ L_{12}^* & L_2^* \end{bmatrix}.$$

We can similarly construct unbounded operators in $\tilde{\mathcal{X}}$. Hence, define the closed-loop dynamics operator $\tilde{A}: \mathcal{D}(\tilde{A}) \subset \tilde{\mathcal{X}} \rightarrow \tilde{\mathcal{X}}$ on the dense domain $\mathcal{D}(\tilde{A}) \triangleq \mathcal{D}(A) \times \mathbb{R}^{n_c}$ by $\tilde{A}\tilde{x} = (Ax + BC_c x_c, B_c Cx + A_c x_c)$. Since \tilde{A} can be represented by

$$\tilde{A} = \begin{bmatrix} A & BC_c \\ B_c C & A_c \end{bmatrix} = \begin{bmatrix} A & 0 \\ 0 & 0 \end{bmatrix} + \begin{bmatrix} 0 & BC_c \\ B_c C & A_c \end{bmatrix}$$

and since the closed-loop operator

$$\begin{bmatrix} A & 0 \\ 0 & 0 \end{bmatrix}: \mathcal{D}(\tilde{A}) \rightarrow \tilde{\mathcal{X}}$$

generates the strongly continuous semigroup

$$\begin{bmatrix} e^{At} & 0 \\ 0 & I_{n_c} \end{bmatrix}, \quad t \geq 0,$$

it follows from [50, Thm., p. 497] that \tilde{A} is also closed and generates a strongly continuous semigroup $e^{\tilde{A}t} \in \mathcal{B}(\tilde{\mathcal{X}})$, $t \geq 0$. To guarantee that J is finite and independent

³ This assumption and its analogue, the lack of a cross-weighting term $x(t)^T R_{12} u(t)$ in (3.5), are for convenience only. See § 5.

⁴ We must require that R_1 and V_1 be nuclear since covariance operators in the white noise formulation of [6] are not necessarily trace class as they are in the formulation of [4].

of initial conditions we restrict our attention to the set of admissible stabilizing compensators

$$\mathcal{A} \triangleq \{(A_c, B_c, C_c): e^{\hat{A}_c t} \text{ is exponentially stable}\}.$$

Hence if $(A_c, B_c, C_c) \in \mathcal{A}$ then there exist $\alpha > 0$ and $\beta > 0$ such that

$$(3.6) \quad \|e^{\hat{A}_c t}\| \leq \alpha e^{-\beta t}, \quad t \geq 0.$$

Since the value of J is independent of the internal realization of the compensator, we can further restrict our attention to

$$\mathcal{A}_+ \triangleq \{(A_c, B_c, C_c) \in \mathcal{A}: (A_c, B_c) \text{ is controllable and } (C_c, A_c) \text{ is observable}\}.$$

The following lemma is required for the statement of the Main Theorem.

LEMMA 3.1. Suppose $\hat{Q}, \hat{P} \in \mathcal{B}(\mathcal{H})$ have finite rank and are nonnegative definite. Then $\hat{Q}\hat{P}$ is nonnegative semisimple. Furthermore, if $\rho(\hat{Q}\hat{P}) = n_c$ then there exist $G, \Gamma \in \mathcal{B}(\mathcal{H}, \mathbb{R}^{n_c})$ and positive-semisimple $M \in \mathbb{R}^{n_c \times n_c}$ such that

$$(3.7) \quad \hat{Q}\hat{P} = G^* M \Gamma,$$

$$(3.8) \quad \Gamma G^* = I_{n_c}.$$

Proof. By Lemma 2.4 there exists a finite-dimensional subspace $\mathcal{M} \subset \mathcal{H}$ such that $\hat{Q}\mathcal{M} \subset \mathcal{M}$, $\hat{Q}\mathcal{M}^\perp = 0$, $\hat{P}\mathcal{M} \subset \mathcal{M}$ and $\hat{P}\mathcal{M}^\perp = 0$. Hence there exists an orthonormal basis for \mathcal{H} with respect to which \hat{Q} and \hat{P} have the infinite-matrix representations

$$\hat{Q} = \begin{bmatrix} \hat{Q}_1 & 0 \\ 0 & 0 \end{bmatrix}, \quad \hat{P} = \begin{bmatrix} \hat{P}_1 & 0 \\ 0 & 0 \end{bmatrix},$$

where $\hat{Q}_1, \hat{P}_1 \in \mathbb{R}^{r \times r}$ are nonnegative definite and $r \triangleq \dim \mathcal{M}$. Since by Lemma 2.6 there exists invertible $\Psi \in \mathbb{R}^{r \times r}$ such that $\tilde{\Lambda} = \Psi^{-1} \hat{Q}_1 \hat{P}_1 \Psi$ is nonnegative diagonal, we have

$$\hat{Q}\hat{P} = \begin{bmatrix} \Psi & 0 \\ 0 & I_\infty \end{bmatrix} \begin{bmatrix} \tilde{\Lambda} & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \Psi^{-1} & 0 \\ 0 & I_\infty \end{bmatrix},$$

which shows that $\hat{Q}\hat{P}$ is nonnegative semisimple. If, furthermore, $\rho(\hat{Q}\hat{P}) = n_c$ then it is clear that Ψ can be chosen (i.e., modified by an orthogonal matrix) so that

$$\tilde{\Lambda} = \begin{bmatrix} \Lambda & 0 \\ 0 & 0 \end{bmatrix},$$

where $\Lambda \in \mathbb{R}^{n_c \times n_c}$ is positive diagonal. Hence,

$$\hat{Q}\hat{P} = \begin{bmatrix} \Psi & 0 \\ 0 & I_\infty \end{bmatrix} \begin{bmatrix} I_{n_c} \\ 0 \\ 0 \end{bmatrix} \Lambda \begin{bmatrix} I_{n_c} & 0 & 0 \end{bmatrix} \begin{bmatrix} \Psi^{-1} & 0 \\ 0 & I_\infty \end{bmatrix},$$

which shows that (3.7) and (3.8) are satisfied with

$$G = \begin{bmatrix} [S^T & 0] & 0 \end{bmatrix} \begin{bmatrix} \Psi^T & 0 \\ 0 & I_\infty \end{bmatrix}, \quad M = S^{-1} \Lambda S, \quad \Gamma = \begin{bmatrix} [S^{-1} & 0] & 0 \end{bmatrix} \begin{bmatrix} \Psi^{-1} & 0 \\ 0 & I_\infty \end{bmatrix},$$

for all invertible $S \in \mathbb{R}^{n_c \times n_c}$. \square

We shall refer to $G, \Gamma \in \mathcal{B}(\mathcal{H}, \mathbb{R}^{n_c})$ and positive-semisimple $M \in \mathbb{R}^{n_c \times n_c}$ satisfying (3.7) and (3.8) as a (G, M, Γ) -factorization of $\hat{Q}\hat{P}$. For convenience in stating the Main Theorem define

$$\Sigma \triangleq B R_2^{-1} B^*, \quad \bar{\Sigma} \triangleq C^* V_2^{-1} C.$$

MAIN THEOREM. Suppose $(A_\alpha, B_\alpha, C_\alpha) \in \mathcal{A}_+$ solves the steady-state fixed-order dynamic-compensation problem. Then there exist nonnegative-definite $Q, P, \hat{Q}, \hat{P} \in \mathcal{B}_1(\mathcal{H})$ such that A_c, B_c and C_c are given by

$$(3.9) \quad A_c = \Gamma(A - Q\bar{\Sigma} - \Sigma P)G^*,$$

$$(3.10) \quad B_c = \Gamma Q C^* V_2^{-1},$$

$$(3.11) \quad C_c = -R_2^{-1} B^* P G^*,$$

for some (G, M, Γ) -factorization of $\hat{Q}\hat{P}$, and such that, with $\tau \triangleq G^*\Gamma$, the following conditions are satisfied:

$$(3.12a, b) \quad Q: \mathcal{D}(A^*) \rightarrow \mathcal{D}(A), \quad P: \mathcal{D}(A) \rightarrow \mathcal{D}(A^*),$$

$$(3.13a, b) \quad \hat{Q}: \mathcal{H} \rightarrow \mathcal{D}(A), \quad \hat{P}: \mathcal{H} \rightarrow \mathcal{D}(A^*),$$

$$(3.14a, b, c)^5 \quad \rho(\hat{Q}) = \rho(\hat{P}) = \rho(\hat{Q}\hat{P}) = n_\alpha,$$

$$(3.15) \quad 0 = (A - \tau Q\bar{\Sigma})Q + Q(A - \tau Q\bar{\Sigma})^* + V_1 + \tau Q\bar{\Sigma} Q \tau^*,$$

$$(3.16) \quad 0 = (A - \Sigma P\tau)^* P + P(A - \Sigma P\tau) + R_1 + \tau^* P \Sigma P \tau,$$

$$(3.17) \quad 0 = [(A - \Sigma P)\hat{Q} + \hat{Q}(A - \Sigma P)^* + Q\bar{\Sigma}Q]\tau^*,$$

$$(3.18) \quad 0 = [(A - Q\bar{\Sigma})^* \hat{P} + \hat{P}(A - Q\bar{\Sigma}) + P\Sigma P]\tau.$$

The content of the Main Theorem is clearly a set of necessary conditions which characterize the optimal steady-state fixed-order dynamic compensator when it exists. These necessary conditions consist of a system of four operator equations including a pair of modified Riccati equations (3.15) and (3.16) and a pair of modified Lyapunov equations (3.17) and (3.18). The salient feature of these four equations is the coupling by the operator $\tau \in \mathcal{B}(\mathcal{H})$ which, because of (3.8), is idempotent, i.e., $\tau^2 = \tau$. In general, τ is an oblique projection and not an orthogonal projection since there is no requirement that τ be selfadjoint. Additional features of the Main Theorem will be discussed in the remainder of this section. For convenience, let $G, M, \Gamma, \tau, Q, P, \hat{Q}$ and \hat{P} be as given by the Main Theorem and define $\Lambda \triangleq \text{diag}(\lambda_1, \dots, \lambda_n)$, where $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_n > 0$ are the eigenvalues of M .

We begin by noting that if x_c is replaced by Sx_c where $S \in \mathbb{R}^{n \times n}$ is invertible, then an "equivalent" compensator is obtained with $(A_\alpha, B_\alpha, C_\alpha)$ replaced by $(SA_\alpha S^{-1}, SB_\alpha, C_\alpha S^{-1})$.

PROPOSITION 3.1. Let $(A_\alpha, B_\alpha, C_\alpha) \in \mathcal{A}_+$. If $S \in \mathbb{R}^{n \times n}$ is invertible then $(SA_\alpha S^{-1}, SB_\alpha, C_\alpha S^{-1}) \in \mathcal{A}_+$ and

$$(3.19) \quad J(A_\alpha, B_\alpha, C_\alpha) = J(SA_\alpha S^{-1}, SB_\alpha, C_\alpha S^{-1}).$$

Proof. Although the result is obvious from system-theoretic arguments, we shall prove it analytically by utilizing elements of the development in § 4. Define

$$\tilde{S} \triangleq \begin{bmatrix} I_\infty & 0 \\ 0 & S \end{bmatrix} \in \mathcal{B}(\tilde{\mathcal{H}})$$

and note that replacing $(A_\alpha, B_\alpha, C_\alpha)$ by $(SA_\alpha S^{-1}, SB_\alpha, C_\alpha S^{-1})$ is equivalent to replacing \tilde{A}, \tilde{V} and \tilde{R} by $\tilde{S}\tilde{A}\tilde{S}^{-1}, \tilde{S}\tilde{V}\tilde{S}^*$ and $\tilde{S}^{-*}\tilde{R}\tilde{S}^{-1}$, respectively. If $\alpha, \beta > 0$ satisfy (3.6) then a straightforward application of the Hille-Yosida theorem [57, pp. 153-5] shows that

⁵(3.14a) refers to $\rho(\hat{Q}) = n_\alpha$ etc.

the strongly continuous semigroup generated by $\tilde{S}\tilde{A}\tilde{S}^{-1}$ satisfies $\|e^{\tilde{S}\tilde{A}\tilde{S}^{-1}t}\| \leq \|\tilde{S}\| \|\tilde{S}^{-1}\| \alpha e^{-\beta t}$, which proves the first assertion. Since $\tilde{S}e^{\tilde{A}t}\tilde{S}^{-1}$, $t \geq 0$, is also a strongly continuous semigroup with generator $\tilde{S}\tilde{A}\tilde{S}^{-1}$, it follows that $\tilde{S}e^{\tilde{A}t}\tilde{S}^{-1} = e^{\tilde{S}\tilde{A}\tilde{S}^{-1}t}$. Hence

$$\int_0^\infty e^{\tilde{S}\tilde{A}\tilde{S}^{-1}t} (\tilde{S}\tilde{V}\tilde{S}^*) e^{(\tilde{S}\tilde{A}\tilde{S}^{-1})^*t} dt = \tilde{S}\tilde{Q}\tilde{S}^*$$

and (3.19) follows from $\text{tr } \tilde{Q}\tilde{R} = \text{tr } (\tilde{S}\tilde{Q}\tilde{S}^*)(\tilde{S}^{-*}\tilde{R}\tilde{S}^{-1})$. \square

In view of Proposition 3.1 one would expect the Main Theorem to apply also to $(SA, S^{-1}, SB, C, S^{-1})$. Indeed, it may be noted that no claim was made as to the uniqueness of the (G, M, Γ) -factorization of $\hat{Q}\hat{P}$ used to determine A, B , and C in (3.9)–(3.11). These observations are reconciled by the following result which shows that a transformation of the compensator state basis corresponds to the alternative factorization $\hat{Q}\hat{P} = (S^{-T}G)^T(SMS^{-1})(S\Gamma)$ and, moreover, that all (G, M, Γ) -factorizations of $\hat{Q}\hat{P}$ are related by a nonsingular transformation. Note that τ remains invariant over the class of factorizations.

PROPOSITION 3.2. *If $S \in \mathbb{R}^{n_c \times n_c}$ is invertible then $\bar{G} \triangleq S^{-T}G$, $\bar{\Gamma} \triangleq S\Gamma$ and $\bar{M} \triangleq SMS^{-1}$ satisfy*

$$(3.7)' \quad \hat{Q}\hat{P} = \bar{G}^* \bar{M} \bar{\Gamma},$$

$$(3.8)' \quad \bar{\Gamma} \bar{G}^* = I_{n_c}.$$

Conversely, if $\bar{G}, \bar{\Gamma} \in \mathcal{B}(\mathcal{H}, \mathbb{R}^{n_c})$ and invertible $\bar{M} \in \mathbb{R}^{n_c \times n_c}$ satisfy (3.7)' and (3.8)', then there exists invertible $S \in \mathbb{R}^{n_c \times n_c}$ such that $\bar{G} = S^{-T}G$, $\bar{\Gamma} = S\Gamma$ and $\bar{M} = SMS^{-1}$.

Proof. The first part of the proposition is immediate. The second part follows by taking $S \triangleq \bar{M}^{-1}\bar{\Gamma}G^*M^{-1}$, noting $S^{-1} = M\bar{\Gamma}\bar{G}^*\bar{M}^{-1}$ and using the identities $\bar{\Gamma}G^*M\bar{\Gamma}\bar{G}^* = \bar{M}$ and $M\bar{\Gamma}\bar{G}^* = \bar{\Gamma}\bar{G}^*\bar{M}$. \square

The next result shows that there exists a similarity transformation which simultaneously diagonalizes $\hat{Q}\hat{P}$ and τ .

PROPOSITION 3.3. *There exists invertible $\Phi \in \mathcal{B}(\mathcal{H})$ such that*

$$(3.20a, b) \quad \hat{Q} = \Phi^{-1} \begin{bmatrix} \Lambda_{\hat{Q}} & 0 \\ 0 & 0 \end{bmatrix} \Phi^{-*}, \quad \hat{P} = \Phi^* \begin{bmatrix} \Lambda_{\hat{P}} & 0 \\ 0 & 0 \end{bmatrix} \Phi,$$

$$(3.21a, b) \quad \hat{Q}\hat{P} = \Phi^{-1} \begin{bmatrix} \Lambda & 0 \\ 0 & 0 \end{bmatrix} \Phi, \quad \tau = \Phi^{-1} \begin{bmatrix} I_{n_c} & 0 \\ 0 & 0 \end{bmatrix} \Phi,$$

where $\Lambda_{\hat{Q}}, \Lambda_{\hat{P}} \in \mathbb{R}^{n_c \times n_c}$ are positive diagonal and $\Lambda_{\hat{Q}}\Lambda_{\hat{P}} = \Lambda$. Consequently,

$$(3.22a, b) \quad \hat{Q} = \tau\hat{Q}, \quad \hat{P} = \hat{P}\tau.$$

Proof. Proceeding as in the proof of Lemma 3.1, choose an orthonormal basis for \mathcal{H} with respect to which

$$\hat{Q} = \begin{bmatrix} \hat{Q}_1 & 0 \\ 0 & 0 \end{bmatrix} \quad \text{and} \quad \hat{P} = \begin{bmatrix} \hat{P}_1 & 0 \\ 0 & 0 \end{bmatrix},$$

where $\hat{Q}_1, \hat{P}_1 \in \mathbb{R}^{r \times r}$ are nonnegative definite. By [55, Thm. 6.2.5, p. 123], there exists invertible $\Psi \in \mathbb{R}^{r \times r}$ such that $\tilde{\Lambda}_{\hat{Q}} \triangleq \Psi\hat{Q}_1\Psi^T$ and $\tilde{\Lambda}_{\hat{P}} = \Psi^{-T}\hat{P}_1\Psi^{-1}$ are nonnegative diagonal. Because of (3.14), it is clear that Ψ can be chosen so that

$$\tilde{\Lambda}_{\hat{Q}} = \begin{bmatrix} \Lambda_{\hat{Q}} & 0 \\ 0 & 0 \end{bmatrix} \quad \text{and} \quad \tilde{\Lambda}_{\hat{P}} = \begin{bmatrix} \Lambda_{\hat{P}} & 0 \\ 0 & 0 \end{bmatrix},$$

where $\Lambda_Q, \Lambda_P \in \mathbb{R}^{n_c \times n_c}$ are positive diagonal. Thus (3.20) holds with

$$\Phi \triangleq \begin{bmatrix} \Psi & 0 \\ 0 & I_\infty \end{bmatrix}.$$

From (3.20) it follows that

$$\hat{Q}\hat{P} = \Phi^{-1} \begin{bmatrix} \Lambda_Q \Lambda_P & 0 \\ 0 & 0 \end{bmatrix} \Phi.$$

Now define $\bar{G} = [I_{n_c} \ 0] \Phi^{-*}$, $\bar{M} = \Lambda_Q \Lambda_P$ and $\bar{\Gamma} = [I_{n_c} \ 0] \Phi$ so that (3.7)' and (3.8)' are satisfied. By the second part of Proposition 3.2 there exists invertible $S \in \mathbb{R}^{n_c \times n_c}$ such that $G = S^T \bar{G}$, $M = S^{-1} \bar{M} S$ and $\Gamma = S^{-1} \bar{\Gamma}$. Since M and \bar{M} have the same eigenvalues, $\bar{M} = \Lambda$ (modulo an ordering of the diagonal elements) and thus (3.21a) holds. Finally, (3.21b) follows from

$$\tau = G^* \Gamma = \bar{G}^* \bar{\Gamma} = \Phi^{-1} \begin{bmatrix} I_{n_c} & 0 \\ 0 & 0 \end{bmatrix} \Phi. \quad \square$$

Remark 3.1. Proposition 3.3 shows that $\lambda_1, \dots, \lambda_{n_c}$ are the positive eigenvalues of $\hat{Q}\hat{P}$.

Remark 3.2. The simultaneous diagonalization in (3.20) has been effected by a *contragredient transformation* [55], [58]. For applications of this type of transformation to model reduction and realization problems see [12], [59]–[61]. Simultaneous diagonalization of operators is discussed in [53, p. 181].

The following result validates the precise handling of the unbounded operator A in (3.9), (3.17) and (3.18).

PROPOSITION 3.4. *The following relations hold:*

$$(3.23a, b, c) \quad \rho(G) = \rho(\Gamma) = \rho(\tau) = n_c$$

$$(3.24a, b) \quad \tau: \mathcal{H} \rightarrow \mathcal{D}(A), \quad \tau^*: \mathcal{H} \rightarrow \mathcal{D}(A^*),$$

$$(3.25a, b) \quad G^*: \mathbb{R}^{n_c} \rightarrow \mathcal{D}(A), \quad \Gamma^*: \mathbb{R}^{n_c} \rightarrow \mathcal{D}(A^*).$$

Proof. From (3.8) and (2.1) it follows that $n_c = \rho(\Gamma G^*) \leq \min \{\rho(\Gamma), \rho(G^*)\}$. Since $\rho(\Gamma) \leq n_c$, $\rho(G) = \rho(G^*)$ and $\rho(G) \leq n_c$, (3.23a) and (3.23b) hold. To show (3.23c) either note (3.21b) or use (3.14a) and (3.22) to obtain

$$n_c = \rho(\hat{Q}) = \rho(\tau \hat{Q}) \leq \rho(\tau) = \rho(G^* \Gamma) \leq \rho(\Gamma) = n_c$$

To prove (3.24a) note that (3.22a) implies $\mathcal{R}(\hat{Q}) \subset \mathcal{R}(\tau)$ and thus $\rho(\hat{Q}) = \rho(\tau)$ implies $\mathcal{R}(\hat{Q}) = \mathcal{R}(\tau)$, and similarly for (3.24b). Finally, (3.25) follows from (3.23), (3.24), the definition $\tau = G^* \Gamma$ and the fact that $\tau^* = \Gamma^* G$. \square

Since the domain of A may not be all of \mathcal{H} , expressions involving A require special interpretation. First note that because of the range condition (3.25a), the expression (3.9) indeed represents an $n_c \times n_c$ matrix (see, e.g., [6, p. 80]). Similarly, because of (3.25b), A_c^T is given by

$$(3.26) \quad A_c^T = G(A^* - \Sigma Q - P \Sigma) \Gamma^*.$$

With regard to (3.15), note that because of (3.12a), the right-hand side of (3.15) is a linear operator with domain $\mathcal{D}(A^*)$. Since $\Theta \triangleq -\tau Q \Sigma Q - Q \Sigma Q \tau^* + V_1 + \tau Q \Sigma Q \tau^*$ is continuous on $\mathcal{D}(A^*)$, $AQ + QA^*$ has a continuous extension on \mathcal{H} given precisely by $-\Theta$. Similar remarks apply to (3.16). Analogous domain conditions were obtained in [5] for a deterministic infinite-dimensional linear-quadratic control problem with

full-state feedback. Finally, because of (3.24) the right-hand sides of (3.17) and (3.18) denote bounded linear operators on all of \mathcal{X} .

It is useful to present an alternative form of the optimal projection equations (3.15)–(3.18). For convenience define the notation

$$\tau_{\perp} \triangleq I_{\mathcal{X}} - \tau.$$

PROPOSITION 3.5. *Equations (3.15)–(3.18) are equivalent, respectively, to*

$$(3.27) \quad 0 = AQ + QA^* + V_1 - Q\bar{\Sigma}Q + \tau_{\perp}Q\bar{\Sigma}Q\tau_{\perp}^*,$$

$$(3.28) \quad 0 = A^*P + PA + R_1 - P\Sigma P + \tau_{\perp}^*P\Sigma P\tau_{\perp},$$

$$(3.29) \quad 0 = (A - \Sigma P)\hat{Q} + \hat{Q}(A - \Sigma P)^* + Q\bar{\Sigma}Q - \tau_{\perp}Q\bar{\Sigma}Q\tau_{\perp}^*,$$

$$(3.30) \quad 0 = (A - Q\bar{\Sigma})^*\hat{P} + \hat{P}(A - Q\bar{\Sigma}) + P\Sigma P - \tau_{\perp}^*P\Sigma P\tau_{\perp}.$$

Proof. The equivalence of (3.27) and (3.28) to (3.15) and (3.16) is immediate. Using (3.22a) in the form $\hat{Q} = \hat{Q}\tau^*$, we obtain (3.17) = (3.29) τ^* . Conversely, from (3.22a) and $[(A - \Sigma P)\hat{Q}]^* = \hat{Q}(A - \Sigma P)^*$ (see, e.g., [6, p. 80]) it follows that (3.29) = (3.17) + (3.17) * - $\tau(3.17)$. Similarly, (3.18) and (3.30) are equivalent. \square

The form of the optimal projection equations (3.27)–(3.30) helps demonstrate the relationship between the Main Theorem and the classical *LQG* result when $\dim \mathcal{X} = n < \infty$. In this case we need only note that the (G, M, Γ) -factorization of $\hat{Q}\hat{P}$ in the “full-order” case $n_c = n$ is given by $G = \Gamma = I_n$ and $M = \hat{Q}\hat{P}$. Since $\tau = I_n$, and thus $\tau_{\perp} = 0$, (3.27) and (3.28) reduce to the standard observer and regulator Riccati equations and (3.9)–(3.11) yield the usual *LQG* expressions. Furthermore, note that in the full-order case

$$(3.31) \quad A_c = A + BC_c - B_cC$$

and (3.29) and (3.31) can be written as

$$(3.32) \quad 0 = (A_c + B_cC)\hat{Q} + \hat{Q}(A_c + B_cC)^T + B_cV_2B_c^T,$$

$$(3.33) \quad 0 = (A_c - BC_c)^T\hat{P} + \hat{P}(A_c - BC_c) + C_c^TR_2C_c$$

Since, as is well known, the stability of \tilde{A} corresponds to the stability of $A + BC_c = A_c + B_cC$ and $A - B_cC = A_c - BC_c$, it follows from standard results (e.g., [62, pp. 48, 277]) that the positive-definiteness conditions (3.14a, b) are equivalent to the assumption that (A_c, B_c, C_c) is controllable and observable.

To obtain a geometric interpretation of the optimal projection we introduce the quasi-full-state estimate

$$\hat{x}(t) \triangleq G^*x_c(t) \in \mathcal{X}$$

so that $\tau\hat{x}(t) = \hat{x}(t)$ and $x_c(t) = \Gamma\hat{x}(t)$. Now, the closed-loop system (3.1)–(3.4) can be written as

$$(3.34) \quad \dot{x}(t) = Ax(t) - B\hat{C}_c\tau\hat{x}(t) + H_1w(t),$$

$$(3.35) \quad \dot{\hat{x}}(t) = \tau(A + B\hat{C}_c - \hat{B}_cC)\tau\hat{x}(t) + \tau\hat{B}_c(Cx(t) + H_2w(t)),$$

where (3.35) is interpreted in the sense of (3.34) since $\hat{x}(t) \in \mathcal{X}$ and where

$$\hat{B}_c \triangleq QC^*V_2^{-1}, \quad \hat{C}_c \triangleq -R_2^{-1}B^*P.$$

It can thus be seen that the geometric structure of the quasi-full-order compensator is entirely dictated by the projection τ . In particular, control inputs $\tau\hat{x}(t)$ determined by

(3.35) are contained in $\mathcal{R}(\tau)$ and sensor inputs $\tau \hat{B}_c y(t)$ are annihilated unless they are contained in $[\mathcal{N}(\tau)]^\perp = \mathcal{R}(\tau^*)$. Consequently, $\mathcal{R}(\tau)$ and $\mathcal{R}(\tau^*)$ are the control and observation subspaces, respectively, of the compensator. Since τ is not necessarily an orthogonal projection, these (finite-dimensional) subspaces may be different.

From the form of (3.35) it is tempting to suggest that the optimal fixed-order dynamic compensator can be obtained by projecting the full-order (infinite-dimensional) LQG compensator. However, this is generally impossible for the following simple reason. Although the expressions for A_c , B_c and C_c in (3.9)–(3.11) have the form of a projection of the full-order LQG compensator, the operators Q and P in (3.9)–(3.11) are *not* the solutions of the usual LQG Riccati equations but instead must be obtained by simultaneously solving all four coupled equations (3.15)–(3.18). This observation reinforces the statement made in § 1 that the optimal fixed-order dynamic compensator cannot in general be obtained by LQG followed by closed-loop controller reduction as in [14] and [15].

We now give an explicit characterization of the optimal projection in terms of \hat{Q} and \hat{P} . Since $\hat{Q}\hat{P}$ has finite rank, its Drazin inverse $(\hat{Q}\hat{P})^D$ exists (see [63, Thm. 6, p. 108]) and, since $(\hat{Q}\hat{P})^2 = G^* M^2 \Gamma$, and hence $\rho(\hat{Q}\hat{P})^2 = \rho(\hat{Q}\hat{P})$, the "index" of $\hat{Q}\hat{P}$ (see [63], [64]) is 1. In this case the Drazin inverse is traditionally called the group inverse and is denoted by $(\hat{Q}\hat{P})^\#$ (see, e.g., [64, p. 124] or [65]).

PROPOSITION 3.6. *The optimal projection τ is given by*

$$(3.36) \quad \tau = \hat{Q}\hat{P}(\hat{Q}\hat{P})^\#.$$

Proof. It is easy to verify that the conditions characterizing the Drazin inverse [63] for the case that $\hat{Q}\hat{P}$ has index 1 are satisfied by $G^* M^{-1} \Gamma$. Hence $(\hat{Q}\hat{P})^\# = G^* M^{-1} \Gamma$ and (3.8) implies (3.36). \square

We now give an alternative characterization of the optimal projection by introducing the following notation from [51, p. 73]. For $\phi, \psi \in \mathcal{H}$ define the operator $\phi \otimes \psi \in \mathcal{B}(\mathcal{H})$ by

$$(\phi \otimes \psi)x \triangleq \langle x, \phi \rangle \psi, \quad x \in \mathcal{H},$$

and note that $\rho(\phi \otimes \psi) = 1$ if ϕ and ψ are both nonzero and $(\phi \otimes \psi)^* = \psi \otimes \phi$. Using this notation, (3.21a) can be written as

$$(3.37) \quad \Phi \hat{Q}\hat{P}\Phi^{-1} = \sum_{i=1}^n \lambda_i \xi_i \otimes \xi_i,$$

where $\{\xi_i\}_{i=1}^\infty$ is an orthonormal basis for \mathcal{H} . In terms of the Riesz bases (see e.g., [52, p. 309])

$$\phi_i \triangleq \Phi^* \xi_i, \quad \psi_i \triangleq \Phi^{-1} \xi_i, \quad i = 1, 2, \dots,$$

(3.37) is equivalent to

$$(3.38) \quad \hat{Q}\hat{P} = \sum_{i=1}^n \lambda_i \phi_i \otimes \psi_i,$$

which can be regarded as a specialized spectral decomposition of a semisimple operator. We emphasize that, in contrast to the singular value decomposition for compact nonnormal operators (see, e.g., [50, p. 261]), the λ_i in (3.38) are *eigenvalues* of $\hat{Q}\hat{P}$ (see Remark 3.1), not singular values. Moreover, although $\{\phi_i\}_{i=1}^\infty$ and $\{\psi_i\}_{i=1}^\infty$ are bases for \mathcal{H} , they are not necessarily orthogonal. They are, however, biorthonormal, i.e., $\langle \phi_i, \psi_j \rangle = \delta_{ij}$, and hence $\phi_i \otimes \psi_i$ is a rank-one projection and $(\phi_i \otimes \psi_i)(\phi_j \otimes \psi_j) = 0$, $i \neq j$.

Since τ is a rank- n_c projection, it is not surprising that τ is given precisely by

$$(3.39) \quad \tau = \sum_{i=1}^{n_c} \phi_i \otimes \psi_i.$$

The following result summarizes the above observations.

PROPOSITION 3.7. *There exist biorthonormal linearly independent sets $\{\psi_i\}_{i=1}^{n_c} \subset \mathcal{D}(A)$ and $\{\phi_i\}_{i=1}^{n_c} \subset \mathcal{D}(A^*)$ such that (3.38) and (3.39) hold. Furthermore, if the (G, M, Γ) -factorization of $\hat{Q}\hat{P}$ is chosen such that $M = \Lambda$, then, for all $x \in \mathcal{X}$,*

$$Gx = (\langle x, \psi_1 \rangle, \dots, \langle x, \psi_{n_c} \rangle)^T,$$

$$\Gamma x = (\langle x, \phi_1 \rangle, \dots, \langle x, \phi_{n_c} \rangle)^T.$$

Remark 3.3. Note that $\hat{P}\hat{Q}$ and τ^* are given by

$$\hat{P}\hat{Q} = \sum_{i=1}^{n_c} \lambda_i \psi_i \otimes \phi_i, \quad \tau^* = \sum_{i=1}^{n_c} \psi_i \otimes \phi_i,$$

and, for all $y \triangleq (y_1, \dots, y_{n_c})^T \in \mathbb{R}^{n_c}$, G^* and Γ^* satisfy

$$G^*y = \sum_{i=1}^{n_c} y_i \psi_i, \quad \Gamma^*y = \sum_{i=1}^{n_c} y_i \phi_i.$$

4. Proof of the Main Theorem. We state and prove a series of lemmas which allow us to compute the Frechet derivatives of J with respect to A_c , B_c and C_c . Requiring that these derivatives vanish leads to the necessary conditions in their "primitive" form. A transformation of variables then leads to the form of the necessary conditions (3.9)–(3.18).

Let "u-lim" denote the uniform limit (i.e., limit in operator norm) for bounded linear operators [50, p. 150] and, for strongly continuous $S(t) \in \mathcal{B}(\mathcal{X})$, $t \geq 0$, interpret the strong integral $\int_0^t S(t) dt$ according to $\int_0^t S(t)z dt$, $z \in \mathcal{X}$ [50, p. 152]. Also recall the standard fact [6, p. 186] that $(e^{A'})^* = e^{A'^*}$ and similarly for \tilde{A} . Throughout this section let $(A_c, B_c, C_c) \in \mathcal{A}_+$ and let $\alpha, \beta > 0$ satisfy (3.6).

To begin, note that the closed-loop system (3.1)–(3.4) can be written as

$$(4.1) \quad \dot{\hat{x}}(t) = \tilde{A}\hat{x}(t) + \tilde{H}w(t),$$

where

$$\tilde{H} \triangleq \begin{bmatrix} H_1 \\ B_c H_2 \end{bmatrix} \in \mathcal{B}_2(\mathcal{X}' \oplus \mathbb{R}^l).$$

For convenience define the nonnegative-definite operator

$$\tilde{V} \triangleq \tilde{H}\tilde{H}^* = \begin{bmatrix} V_1 & 0 \\ 0 & B_c V_2 B_c^T \end{bmatrix} \in \mathcal{B}_1(\tilde{\mathcal{X}}).$$

In terms of the augmented state $\hat{x}(t)$, the performance criterion (3.5) becomes

$$(4.2) \quad J(A_c, B_c, C_c) = \lim_{t \rightarrow \infty} E\langle \tilde{R}\hat{x}(t), \hat{x}(t) \rangle,$$

where the nonnegative-definite operator \tilde{R} is defined by

$$\tilde{R} \triangleq \begin{bmatrix} R_1 & 0 \\ 0 & C_c^T R_2 C_c \end{bmatrix} \in \mathcal{B}_1(\tilde{\mathcal{X}}).$$

To write (4.2) in terms of the covariance of $\tilde{x}(t)$, recall [6, p. 308] that the covariance " $E[(\xi - E\xi)(\xi - E\xi)^*]$ " of a Hilbert-space-valued weak random variable ξ is defined to be the nonnegative-definite operator S which satisfies

$$\langle Sy, z \rangle = E\langle \xi - E\xi, y \rangle \langle \xi - E\xi, z \rangle$$

for all y, z in the Hilbert space. Hence define [6, p. 317]

$$\tilde{Q}(t) \triangleq E[(\tilde{x}(t) - E\tilde{x}(t))(\tilde{x}(t) - E\tilde{x}(t))^*].$$

LEMMA 4.1. $\tilde{Q} \triangleq u\text{-}\lim_{t \rightarrow \infty} \tilde{Q}(t)$ exists and is given by

$$(4.3) \quad \tilde{Q} = \int_0^\infty e^{\tilde{A}t} \tilde{V} e^{\tilde{A}^*t} dt.$$

Furthermore,

$$(4.4) \quad J(A_\alpha, B_\alpha, C_\alpha) = \text{tr } \tilde{Q} \tilde{R}.$$

Proof. First compute (as in [6, p. 317])

$$\begin{aligned} \langle \tilde{Q}(t) \tilde{y}, \tilde{z} \rangle &= E\langle \tilde{x}(t) - e^{\tilde{A}t} E\tilde{x}(0), \tilde{y} \rangle \langle \tilde{x}(t) - e^{\tilde{A}t} E\tilde{x}(0), \tilde{z} \rangle \\ &= E\left\langle \int_0^t e^{\tilde{A}(t-s)} \tilde{H} \tilde{w}(s) ds, \tilde{y} \right\rangle \left\langle \int_0^t e^{\tilde{A}(t-\sigma)} \tilde{H} \tilde{w}(\sigma) d\sigma, \tilde{z} \right\rangle \\ &\quad + \langle \tilde{Q}(0) e^{\tilde{A}^*t} \tilde{y}, e^{\tilde{A}t} \tilde{z} \rangle \\ &= E \int_0^t \int_0^t \langle \tilde{w}(s), \tilde{H}^* e^{\tilde{A}^*(t-s)} \tilde{y} \rangle \langle \tilde{w}(\sigma), \tilde{H}^* e^{\tilde{A}^*(t-\sigma)} \tilde{z} \rangle ds d\sigma \\ &\quad + \langle e^{\tilde{A}t} \tilde{Q}(0) e^{\tilde{A}^*t} \tilde{y}, \tilde{z} \rangle \\ &= \int_0^t \langle e^{\tilde{A}(t-s)} \tilde{V} e^{\tilde{A}^*(t-s)} \tilde{y}, \tilde{z} \rangle ds + \langle e^{\tilde{A}t} \tilde{Q}(0) e^{\tilde{A}^*t} \tilde{y}, \tilde{z} \rangle, \end{aligned}$$

which shows that $\tilde{Q}(t)$ is given by

$$\tilde{Q}(t) = e^{\tilde{A}t} \tilde{Q}(0) e^{\tilde{A}^*t} + \int_0^t e^{\tilde{A}s} \tilde{V} e^{\tilde{A}^*s} ds.$$

Clearly, (4.3) makes sense as a strong integral since

$$\|\tilde{Q}\| \leq \int_0^\infty \|e^{\tilde{A}t} \tilde{V} e^{\tilde{A}^*t}\| dt \leq \alpha^2 \|\tilde{V}\| \int_0^\infty e^{-2\beta t} dt < \infty.$$

To demonstrate uniform convergence it need only be noted that

$$\begin{aligned} \|\tilde{Q} - \tilde{Q}(t)\| &= \sup_{\|\tilde{y}\|=1} \|(\tilde{Q} - \tilde{Q}(t))\tilde{y}\| \\ &= \sup_{\|\tilde{y}\|=1} \left\| \int_t^\infty e^{\tilde{A}s} \tilde{V} e^{\tilde{A}^*s} \tilde{y} ds - e^{\tilde{A}t} \tilde{Q}(0) e^{\tilde{A}^*t} \tilde{y} \right\| \\ &\leq \int_t^\infty \|e^{\tilde{A}s} \tilde{V} e^{\tilde{A}^*s}\| ds + \|e^{\tilde{A}t} \tilde{Q}(0) e^{\tilde{A}^*t}\| \\ &\leq \frac{1}{2} \alpha^2 \|\tilde{V}\| \beta^{-1} e^{-2\beta t} + \|\tilde{Q}(0)\| e^{-2\beta t}. \end{aligned}$$

Next, let $\{\phi_i\}_{i=1}^{\infty}$ be an orthonormal basis for \mathcal{H} and use Parseval's equality to obtain

$$J(A_n, B_n, C_n) = \lim_{t \rightarrow \infty} \mathbb{E} \|\tilde{R}^{1/2} \tilde{x}(t)\|^2 = \lim_{t \rightarrow \infty} \mathbb{E} \sum_{i=1}^{\infty} \langle \tilde{R}^{1/2} \tilde{x}(t), \phi_i \rangle^2.$$

Since

$$f_n(t) \triangleq \sum_{i=1}^n \langle \tilde{R}^{1/2} \tilde{x}(t), \phi_i \rangle^2, \quad t \geq 0,$$

is nonnegative for each n and is increasing in n for each t with limit $\langle \tilde{R} \tilde{x}(t), \tilde{x}(t) \rangle$, monotone convergence permits expectation-limit interchange. Hence using $\mathbb{E} \tilde{x}(t) = e^{\tilde{A}t} \mathbb{E} \tilde{x}(0)$ we have

$$\begin{aligned} J(A_n, B_n, C_n) &= \lim_{t \rightarrow \infty} \sum_{i=1}^{\infty} \mathbb{E} \langle \tilde{x}(t), \tilde{R}^{1/2} \phi_i \rangle^2 \\ &= \lim_{t \rightarrow \infty} \sum_{i=1}^{\infty} [\langle \tilde{Q}(t) \tilde{R}^{1/2} \phi_i, \tilde{R}^{1/2} \phi_i \rangle + \langle e^{\tilde{A}t} \mathbb{E} \tilde{x}(0), \tilde{R}^{1/2} \phi_i \rangle^2] \\ &= \lim_{t \rightarrow \infty} \{ \text{tr} [\tilde{R}^{1/2} \tilde{Q}(t) \tilde{R}^{1/2}] + \| \tilde{R}^{1/2} e^{\tilde{A}t} \mathbb{E} \tilde{x}(0) \|^2 \} \end{aligned}$$

which by Corollary 2.1 yields (4.4). \square

We shall also require the "dual" of \tilde{Q} given by

$$(4.5) \quad \tilde{P} = \int_0^{\infty} e^{\tilde{A}^* t} \tilde{R} e^{\tilde{A} t} dt.$$

Since \tilde{V} and \tilde{R} are nonnegative definite it is readily seen that \tilde{Q} and \tilde{P} are also nonnegative definite.

LEMMA 4.2. $\tilde{Q}, \tilde{P} \in \mathcal{B}_1(\mathcal{H})$.

Proof. It suffices to consider \tilde{Q} only since the situation for \tilde{P} is exactly analogous. Since \tilde{Q} is nonnegative definite, Lemma 2.3 can be used. Letting $\{\phi_i\}_{i=1}^{\infty}$ be an orthonormal basis for \mathcal{H} , we have

$$\begin{aligned} \text{tr } \tilde{Q} &= \sum_{i=1}^{\infty} \langle \tilde{Q} \phi_i, \phi_i \rangle = \sum_{i=1}^{\infty} \left\langle \int_0^{\infty} e^{\tilde{A}^* t} \tilde{V} e^{\tilde{A} t} \phi_i dt, \phi_i \right\rangle \\ &= \lim_{n \rightarrow \infty} \int_0^{\infty} \sum_{i=1}^n \langle \tilde{V} e^{\tilde{A}^* t} \phi_i, e^{\tilde{A} t} \phi_i \rangle dt. \end{aligned}$$

Let $f_n(t)$ denote the above integrand. Since \tilde{V} is nonnegative definite, $\{f_n(\cdot)\}$ is a monotonically increasing sequence of nonnegative functions such that $f_n(t) \rightarrow \text{tr } e^{\tilde{A}^* t} \tilde{V} e^{\tilde{A} t}$, $t \geq 0$. Hence, by monotone convergence and Lemma 2.2,

$$\begin{aligned} \text{tr } \tilde{Q} &= \int_0^{\infty} \text{tr} [e^{\tilde{A}^* t} \tilde{V} e^{\tilde{A} t}] dt \\ &= \int_0^{\infty} \|e^{\tilde{A}^* t} \tilde{V} e^{\tilde{A} t}\|_1 dt \leq \alpha^2 \|\tilde{V}\|_1 \int_0^{\infty} e^{-2\beta t} dt < \infty. \end{aligned} \quad \square$$

LEMMA 4.3. With \tilde{Q} and \tilde{P} given by (4.3) and (4.5) it follows that

$$(4.6) \quad \text{tr } \tilde{Q} \tilde{R} = \text{tr } \tilde{V} \tilde{P}.$$

Proof. For any orthonormal basis $\{\phi_i\}_{i=1}^{\infty}$ of \mathcal{H} we have

$$\begin{aligned}\operatorname{tr} \tilde{Q}\tilde{R} &= \operatorname{tr} \tilde{R}\tilde{Q} = \sum_{i=1}^{\infty} \left\langle \tilde{R} \int_0^{\infty} e^{\tilde{A}t} \tilde{V} e^{\tilde{A}^*t} \phi_i dt, \phi_i \right\rangle \\ &= \lim_{n \rightarrow \infty} \int_0^{\infty} \sum_{i=1}^n \langle \tilde{R} e^{\tilde{A}t} \tilde{V} e^{\tilde{A}^*t} \phi_n, \phi_i \rangle dt.\end{aligned}$$

Letting $f_n(t)$ denote the above integrand it follows that $f_n(t) \rightarrow \operatorname{tr} \tilde{R} e^{\tilde{A}t} \tilde{V} e^{\tilde{A}^*t}$, $t \geq 0$, and

$$|f_n(t)| \leq \sum_{i=1}^{\infty} |\langle e^{\tilde{A}t} \tilde{V} e^{\tilde{A}^*t} \phi_n, \tilde{R} \phi_i \rangle| \leq \alpha^2 \|\tilde{V}\| e^{-2\beta t} \sum_{i=1}^{\infty} \|\tilde{R} \phi_i\|.$$

If $\{\phi_i\}_{i=1}^{\infty}$ is chosen to be the set of orthonormal eigenvectors of \tilde{R} then Lemma 2.1 implies $\sum_{i=1}^{\infty} \|\tilde{R} \phi_i\| = \|\tilde{R}\|$, and thus $|f_n(t)|$ is bounded on $[0, \infty)$ by an integrable function. Hence by dominated convergence,

$$\operatorname{tr} \tilde{Q}\tilde{R} = \int_0^{\infty} \operatorname{tr} [\tilde{R} e^{\tilde{A}t} \tilde{V} e^{\tilde{A}^*t}] dt = \int_0^{\infty} \operatorname{tr} [e^{\tilde{A}^*t} \tilde{R} e^{\tilde{A}t} \tilde{V}] dt = \int_0^{\infty} \sum_{i=1}^{\infty} \langle \tilde{V} \phi_n, e^{\tilde{A}^*t} \tilde{R} e^{\tilde{A}t} \phi_i \rangle dt.$$

And again using dominated convergence,

$$\operatorname{tr} \tilde{Q}\tilde{R} = \sum_{i=1}^{\infty} \int_0^{\infty} \langle \tilde{V} \phi_n, e^{\tilde{A}^*t} \tilde{R} e^{\tilde{A}t} \phi_i \rangle dt = \sum_{i=1}^{\infty} \left\langle \tilde{V} \phi_n, \int_0^{\infty} e^{\tilde{A}^*t} \tilde{R} e^{\tilde{A}t} \phi_i dt \right\rangle = \operatorname{tr} \tilde{V}\tilde{P}. \quad \square$$

The next result is important in that it allows us to treat \tilde{Q} and \tilde{P} as solutions of dual algebraic Lyapunov equations. For a similar result involving groups rather than semigroups see [50, pp. 555-557].

LEMMA 4.4. \tilde{Q} is given by (4.3) if and only if $\tilde{Q} \in \mathcal{B}(\mathcal{H})$ satisfies

$$(4.7) \quad \tilde{Q}: \mathcal{D}(\tilde{A}^*) \rightarrow \mathcal{D}(\tilde{A}),$$

$$(4.8) \quad 0 = \tilde{A}\tilde{Q} + \tilde{Q}\tilde{A}^* + \tilde{V},$$

where (4.8) holds in the sense discussed in § 3. Furthermore, \tilde{P} is given by (4.5) if and only if $\tilde{P} \in \mathcal{B}(\mathcal{H})$ satisfies

$$(4.9) \quad \tilde{P}: \mathcal{D}(\tilde{A}) \rightarrow \mathcal{D}(\tilde{A}^*),$$

$$(4.10) \quad 0 = \tilde{A}^*\tilde{P} + \tilde{P}\tilde{A} + \tilde{R}.$$

Proof. We consider \tilde{Q} only. To prove necessity let $t' > 0$. Then for all $t \in [0, t')$ and $\tilde{x} \in \mathcal{D}(\tilde{A}^*)$ we can write

$$\begin{aligned}e^{\tilde{A}t} \tilde{Q} e^{\tilde{A}^*t'} \tilde{x} &= \int_0^{\infty} e^{\tilde{A}(t+s)} \tilde{V} e^{\tilde{A}^*(t'+s)} \tilde{x} ds \\ &= \int_t^{\infty} e^{\tilde{A}\sigma} \tilde{V} e^{\tilde{A}^*\sigma} e^{\tilde{A}^*(t'-t)} \tilde{x} d\sigma.\end{aligned}$$

Hence,

$$(4.11) \quad \frac{d}{dt} e^{\tilde{A}t} \tilde{Q} e^{\tilde{A}^*t'} \tilde{x} = - \int_t^{\infty} e^{\tilde{A}\sigma} \tilde{V} e^{\tilde{A}^*\sigma} e^{\tilde{A}^*(t'-t)} \tilde{A}^* \tilde{x} d\sigma - e^{\tilde{A}t} \tilde{V} e^{\tilde{A}^*t'} \tilde{x},$$

which shows that $e^{\tilde{A}t} \tilde{Q} e^{\tilde{A}^*t'}$ is strongly differentiable with respect to t for all $t \in [0, t')$. In particular, setting $t = 0$ it follows that $\tilde{Q} e^{\tilde{A}^*t'} \tilde{x} \in \mathcal{D}(\tilde{A})$ for all $\tilde{x} \in \mathcal{D}(\tilde{A}^*)$ (see, e.g., [6, p. 173] or [50, p. 485]). Performing the differentiation on the left-hand side of

(4.11) and setting $t = 0$ yields

$$(4.12) \quad \tilde{A}\tilde{Q}e^{\tilde{A}^*t}\tilde{x} = -\int_0^\infty e^{\tilde{A}\sigma}\tilde{V}e^{\tilde{A}^*\sigma}e^{\tilde{A}^*t}\tilde{A}^*\tilde{x}d\sigma - \tilde{V}e^{\tilde{A}^*t}\tilde{x}.$$

Now fix $\tilde{x} \in \mathcal{D}(\tilde{A}^*)$. Then for $\{t_i\}_{i=1}^\infty$, $t_i > 0$, $t_i \rightarrow 0$, we have

$$\begin{aligned} \tilde{Q}e^{\tilde{A}^*t_i}\tilde{x} &\in \mathcal{D}(\tilde{A}), \quad i = 1, 2, 3, \dots, \\ \tilde{Q}e^{\tilde{A}^*t_i}\tilde{x} &\xrightarrow{t_i \rightarrow \infty} \tilde{Q}\tilde{x}. \end{aligned}$$

Now consider the sequence $\{\tilde{A}\tilde{Q}e^{\tilde{A}^*t_i}\tilde{x}\}_{i=1}^\infty$. Letting $t' = t_i$ in (4.12) and using dominated convergence to interchange limit and integration ($\tilde{A}^*\tilde{x}$ is a fixed element of \mathcal{H}), it follows that

$$(4.13) \quad \lim_{t \rightarrow \infty} \tilde{A}\tilde{Q}e^{\tilde{A}^*t}\tilde{x} = -\int_0^\infty e^{\tilde{A}\sigma}\tilde{V}e^{\tilde{A}^*\sigma}\tilde{A}^*\tilde{x}d\sigma - \tilde{V}\tilde{x}.$$

Since \tilde{A} is closed, $\tilde{Q}\tilde{x} \in \mathcal{D}(\tilde{A})$. This proves (4.7). Also, since \tilde{A} is closed we have

$$\lim_{t \rightarrow \infty} \tilde{A}\tilde{Q}e^{\tilde{A}^*t}\tilde{x} = \tilde{A}\tilde{Q}\tilde{x},$$

which with (4.13) implies

$$\tilde{A}\tilde{Q}\tilde{x} = -\tilde{Q}\tilde{A}^*\tilde{x} - \tilde{V}\tilde{x},$$

and hence

$$(\tilde{A}\tilde{Q} + \tilde{Q}\tilde{A}^* + \tilde{V})\tilde{x} = 0, \quad \tilde{x} \in \mathcal{D}(\tilde{A}^*),$$

as desired.

To prove sufficiency let $\tilde{x} \in \mathcal{D}(\tilde{A})$. Then $e^{\tilde{A}^*t}\tilde{x} \in \mathcal{D}(\tilde{A}^*)$, $t \geq 0$, and hence

$$\frac{d}{dt} e^{\tilde{A}^*t}\tilde{Q}e^{\tilde{A}^*t}\tilde{x} = e^{\tilde{A}^*t}(\tilde{A}\tilde{Q} + \tilde{Q}\tilde{A}^*)e^{\tilde{A}^*t}\tilde{x}.$$

Thus

$$e^{\tilde{A}^*t}\tilde{Q}e^{\tilde{A}^*t}\tilde{x} - \tilde{Q}\tilde{x} = \int_0^t e^{\tilde{A}^*s}(\tilde{A}\tilde{Q} + \tilde{Q}\tilde{A}^*)e^{\tilde{A}^*s}\tilde{x}ds, \quad \tilde{x} \in \mathcal{D}(\tilde{A}^*).$$

Extending $\tilde{A}\tilde{Q} + \tilde{Q}\tilde{A}^*$ to all of \mathcal{H} we obtain

$$e^{\tilde{A}^*t}\tilde{Q}e^{\tilde{A}^*t}\tilde{x} - \tilde{Q}\tilde{x} = -\int_0^t e^{\tilde{A}^*s}\tilde{V}e^{\tilde{A}^*s}\tilde{x}ds, \quad \tilde{x} \in \mathcal{H}.$$

Letting $t \rightarrow \infty$ yields (4.3). \square

We now introduce some notation which will prove to be most convenient in the following results. For $(A'_c, B'_c, C'_c) \in \mathbb{R}^{n_c \times n_c} \times \mathbb{R}^{n_c \times l} \times \mathbb{R}^{m \times n_c}$, define

$$\delta_{A_c} \triangleq A'_c - A_c, \quad \delta_{B_c} \triangleq B'_c - B_c, \quad \delta_{C_c} \triangleq C'_c - C_c$$

and

$$\|(\delta_{A_c}, \delta_{B_c}, \delta_{C_c})\| \triangleq \|\delta_{A_c}\| + \|\delta_{B_c}\| + \|\delta_{C_c}\|.$$

Furthermore, let \tilde{A}' , \tilde{V}' and \tilde{R}' denote \tilde{A} , \tilde{V} and \tilde{R} with (A_c, B_c, C_c) replaced by

(A'_c, B'_c, C'_c) and define

$$\begin{aligned}\delta_{\tilde{A}} \triangleq \tilde{A}' - \tilde{A} &= \begin{bmatrix} 0 & B\delta_{C_c} \\ \delta_{B_c}C & \delta_{A_c} \end{bmatrix}, \\ \delta_{\tilde{V}} \triangleq \tilde{V}' - \tilde{V} &= \begin{bmatrix} 0 & 0 \\ 0 & B_c V_2 \delta_{B_c}^T + \delta_{B_c} V_2 B_c^T + \delta_{B_c} V_2 \delta_{B_c}^T \end{bmatrix}, \\ \delta_{\tilde{R}} \triangleq \tilde{R}' - \tilde{R} &= \begin{bmatrix} 0 & 0 \\ 0 & C_c^T R_2 \delta_{C_c} + \delta_{C_c}^T R_2 C_c + \delta_{C_c}^T R_2 \delta_{C_c} \end{bmatrix}.\end{aligned}$$

We shall also write \tilde{Q}', \tilde{P}' for \tilde{Q}, \tilde{P} as given by (4.3) and (4.5) with $\tilde{A}, \tilde{V}, \tilde{R}$ replaced by $\tilde{A}', \tilde{V}', \tilde{R}'$ and define

$$\delta_{\tilde{Q}} \triangleq \tilde{Q}' - \tilde{Q}, \quad \delta_{\tilde{P}} \triangleq \tilde{P}' - \tilde{P}.$$

LEMMA 4.5. \mathcal{A} is open.

Proof. Let $(A_c, B_c, C_c) \in \mathcal{A}$ be arbitrary and consider the open set

$$(4.14) \quad N \triangleq \{(A'_c, B'_c, C'_c) \in \mathbb{R}^{n_c \times n_c} \times \mathbb{R}^{n_c \times l} \times \mathbb{R}^{m \times n_c} : \|(\delta_{A_c}, \delta_{B_c}, \delta_{C_c})\| < \beta/2\alpha\gamma\},$$

where $\gamma \triangleq \max\{1, \|B\|, \|C\|\}$. Then, since $\tilde{A}' = \tilde{A} + \delta_{\tilde{A}}$ and $\delta_{\tilde{A}} \in \mathcal{B}(\tilde{\mathcal{H}})$ it follows from Theorem 2.1, p. 497 of [50], that for all $(A'_c, B'_c, C'_c) \in N$ and $t \geq 0$,

$$\|e^{\tilde{A}'t}\| \leq \alpha e^{(-\beta + \alpha\|\delta_{\tilde{A}}\|)t} \leq \alpha e^{-\beta t/2}.$$

Hence, $N \subset \mathcal{A}$, as desired. \square

LEMMA 4.6. There exists $c > 0$ such that

$$(4.15) \quad \|\delta_{\tilde{Q}}\| \leq c\|(\delta_{A_c}, \delta_{B_c}, \delta_{C_c})\|,$$

$$(4.16) \quad \|\delta_{\tilde{P}}\| \leq c\|(\delta_{A_c}, \delta_{B_c}, \delta_{C_c})\|,$$

for all $(A'_c, B'_c, C'_c) \in N$, where $N \subset \mathcal{A}$ is the open neighborhood of (A_c, B_c, C_c) defined by (4.14).

Proof. We consider (4.15) only. Since $\|e^{\tilde{A}'t}\| \leq \alpha e^{-\beta t/2}$, $t \geq 0$, $(A'_c, B'_c, C'_c) \in N$, it follows that

$$\begin{aligned}\|\delta_{\tilde{Q}}\| &\leq \int_0^\infty \|e^{\tilde{A}'t} \tilde{V}' e^{\tilde{A}^*t} - e^{\tilde{A}t} \tilde{V} e^{\tilde{A}^*t}\| dt \\ &\leq \int_0^\infty \{\|e^{\tilde{A}'t}\| \|\tilde{V}'\| \|e^{\tilde{A}^*t} - e^{\tilde{A}^*t}\| + \|e^{\tilde{A}'t}\| \|\delta_{\tilde{V}}\| \|e^{\tilde{A}^*t}\| + \|e^{\tilde{A}'t} - e^{\tilde{A}t}\| \|\tilde{V}\| \|e^{\tilde{A}^*t}\|\} dt \\ (4.17) \quad &\leq \alpha(\|\tilde{V}\| + \|\delta_{\tilde{V}}\|) \int_0^\infty \|e^{(\tilde{A} + \delta_{\tilde{A}})^*t} - e^{\tilde{A}^*t}\| e^{-\beta t/2} dt \\ &\quad + \alpha^2 \|\delta_{\tilde{V}}\| \int_0^\infty e^{-3\beta t/2} dt + \alpha \|\tilde{V}\| \int_0^\infty \|e^{(\tilde{A} + \delta_{\tilde{A}})t} - e^{\tilde{A}t}\| e^{-\beta t/2} dt \\ &= \alpha(2\|\tilde{V}\| + \|\delta_{\tilde{V}}\|) \int_0^\infty \|e^{(\tilde{A} + \delta_{\tilde{A}})^*t} - e^{\tilde{A}^*t}\| e^{-\beta t/2} dt + \frac{2\alpha^2}{3\beta} \|\delta_{\tilde{V}}\|.\end{aligned}$$

From [50, p. 497], it follows that the perturbed semigroup $e^{(\tilde{A} + \delta_{\tilde{A}})^*t}$ has an expansion

$$e^{(\tilde{A} + \delta_{\tilde{A}})^*t} = e^{\tilde{A}^*t} + \sum_{i=1}^{\infty} U_i(t), \quad t \geq 0,$$

where $U_i(t) \in \mathcal{B}(\mathcal{H})$, $t \geq 0$, satisfy the estimates

$$\|U_i(t)\| \leq \alpha^{i+1} \|\delta_{\lambda}\|^i e^{-\beta t} t^i / i!.$$

Hence, for all $(A'_c, B'_c, C'_c) \in N$,

$$(4.18) \quad \|e^{(\tilde{A} + \delta_{\lambda})t} - e^{\tilde{A}t}\| \leq \sum_{i=1}^{\infty} \|U_i(t)\| \leq \alpha e^{-\beta t} [e^{\alpha \|\delta_{\lambda}\|} - 1].$$

From (4.17), (4.18) and the relations $\|\delta_{\lambda}\| \leq \gamma \|(\delta_{A_c}, \delta_{B_c}, \delta_{C_c})\| < \beta/2\alpha$ and

$$\int_0^{\infty} [e^{\alpha \|\delta_{\lambda}\|t} - 1] e^{-3\beta t/2} dt < \frac{\alpha\gamma}{3\beta^2} \|(\delta_{A_c}, \delta_{B_c}, \delta_{C_c})\|$$

it follows that

$$\begin{aligned} \|\delta_{\tilde{Q}}\| &\leq \frac{2\alpha^3\gamma}{3\beta^2} (2\|\tilde{V}\| + \|\delta_{\tilde{V}}\|) \|(\delta_{A_c}, \delta_{B_c}, \delta_{C_c})\| \\ &\quad + \frac{2\alpha^2}{3\beta} (2\|B_c V_2\| \|\delta_{B_c}\| + \|V_2\| \|\delta_{B_c}\|^2), \end{aligned}$$

which yields (4.15). \square

Since $\tilde{Q}, \tilde{P} \in \mathcal{B}(\mathcal{H})$ we can write

$$\tilde{Q} = \begin{bmatrix} Q_1 & Q_{12} \\ Q_{12}^* & Q_2 \end{bmatrix}, \quad \tilde{P} = \begin{bmatrix} P_1 & P_{12} \\ P_{12}^* & P_2 \end{bmatrix},$$

where $Q_1 \in \mathcal{B}(\mathcal{H})$, $Q_{12} \in \mathcal{B}(\mathbb{R}^n, \mathcal{H})$, $Q_2 \in \mathbb{R}^{n \times n}$, and similarly for P_1 , P_{12} and P_2 . Note that Q_1 , Q_2 , P_1 and P_2 are nonnegative definite. Also, define the notation

$$\tilde{P}\tilde{Q} = \begin{bmatrix} Z_1 & Z_{12} \\ Z_{21} & Z_2 \end{bmatrix},$$

where

$$\begin{aligned} Z_1 &\triangleq P_1 Q_1 + P_{12} Q_{12}^*, & Z_{12} &\triangleq P_1 Q_{12} + P_{12} Q_2, \\ Z_{21} &\triangleq P_{12}^* Q_1 + P_2 Q_{12}^*, & Z_2 &\triangleq P_{12}^* Q_{12} + P_2 Q_2, \end{aligned}$$

and, for $(A'_c, B'_c, C'_c) \in \mathcal{A}$, let

$$\delta_J(\delta_{A_c}, \delta_{B_c}, \delta_{C_c}) \triangleq J(A'_c, B'_c, C'_c) - J(A_c, B_c, C_c).$$

LEMMA 4.7. Let $(A'_c, B'_c, C'_c) \in \mathcal{A}$. Then

$$(4.19) \quad \delta_J(\delta_{A_c}, \delta_{B_c}, \delta_{C_c}) = \mathcal{L}(\delta_{A_c}, \delta_{B_c}, \delta_{C_c}) + o(\|(\delta_{A_c}, \delta_{B_c}, \delta_{C_c})\|),$$

where

$$(4.20) \quad \begin{aligned} \mathcal{L}(\delta_{A_c}, \delta_{B_c}, \delta_{C_c}) &\triangleq 2 \operatorname{tr} [Z_2 \delta_{A_c}] + 2 \operatorname{tr} [(V_2 B_c^T P_2 + C Z_{21}^*) \delta_{B_c}] \\ &\quad + 2 \operatorname{tr} [Q_2 C_c^T R_2 + Z_{12}^* B] \delta_{C_c} \end{aligned}$$

and

$$(4.21) \quad \lim_{(\delta_{A_c}, \delta_{B_c}, \delta_{C_c}) \rightarrow 0} \|(\delta_{A_c}, \delta_{B_c}, \delta_{C_c})\|^{-1} o(\|(\delta_{A_c}, \delta_{B_c}, \delta_{C_c})\|) = 0.$$

Proof. Combining (4.8) and (4.10) with (4.6), J can be written as

$$J(A_c, B_c, C_c) = \operatorname{tr} [\tilde{Q} \tilde{R} + \tilde{P} \tilde{V}] + \frac{1}{2} \operatorname{tr} [\tilde{Q} \operatorname{cl} (\tilde{A}^* \tilde{P} + \tilde{P} \tilde{A}) + \tilde{P} \operatorname{cl} (\tilde{A} \tilde{Q} + \tilde{Q} \tilde{A}^*)],$$

and likewise for (A'_c, B'_c, C'_c) , where "cl" denotes closure (i.e., extension) of a bounded operator to all of \mathcal{H} . Now using the identity

$$\text{tr}[\tilde{Q}'\tilde{R}' + \tilde{P}'\tilde{V}'] - \text{tr}[\tilde{Q}\tilde{R} + \tilde{P}\tilde{V}] = \text{tr}[\tilde{Q}\delta_{\tilde{R}} + \tilde{P}\delta_{\tilde{V}}] + \text{tr}[\delta_{\tilde{Q}}\tilde{R}' + \delta_{\tilde{P}}\tilde{V}']$$

we can compute

$$\begin{aligned} \delta_J(\delta_{A_c}, \delta_{B_c}, \delta_{C_c}) &= \text{tr}[\tilde{Q}\delta_{\tilde{R}} + \tilde{P}\delta_{\tilde{V}}] + \frac{1}{2} \text{tr}[\tilde{Q} \text{cl}(\tilde{A}^*(\tilde{P} + \delta_{\tilde{P}}) + (\tilde{P} + \delta_{\tilde{P}})\tilde{A}')] \\ &\quad + \frac{1}{2} \text{tr}[\delta_{\tilde{Q}} \text{cl}(\tilde{A}'^*\tilde{P}' + \tilde{P}'\tilde{A}')] \\ &\quad + \frac{1}{2} \text{tr}[\tilde{P} \text{cl}(\tilde{A}'(\tilde{Q} + \delta_{\tilde{Q}}) + (\tilde{Q} + \delta_{\tilde{Q}})\tilde{A}^*)] \\ &\quad + \frac{1}{2} \text{tr}[\delta_{\tilde{P}} \text{cl}(\tilde{A}'\tilde{Q}' + \tilde{Q}'\tilde{A}^*)] \\ &\quad - \frac{1}{2} \text{tr}[\tilde{Q} \text{cl}(\tilde{A}^*\tilde{P} + \tilde{P}\tilde{A}) + \tilde{P} \text{cl}(\tilde{A}\tilde{Q} + \tilde{Q}\tilde{A}^*)] \\ &\quad + \text{tr}[\delta_{\tilde{Q}}\tilde{R}' + \delta_{\tilde{P}}\tilde{V}']. \end{aligned}$$

Using $\tilde{A}' = \tilde{A} + \delta_{\tilde{A}}$ and combining the second, fourth and sixth terms yields

$$\delta_J(\delta_{A_c}, \delta_{B_c}, \delta_{C_c}) = \Lambda + \Omega,$$

where

$$\begin{aligned} \Lambda &\triangleq \text{tr}[\tilde{Q}\delta_{\tilde{R}} + \tilde{P}\delta_{\tilde{V}}] + \frac{1}{2} \text{tr}[\tilde{Q}(\delta_{\tilde{A}}^*\tilde{P} + \tilde{P}\delta_{\tilde{A}}) + \tilde{P}(\delta_{\tilde{A}}\tilde{Q} + \tilde{Q}\delta_{\tilde{A}}^*)] \\ &= \text{tr}[\tilde{Q}\delta_{\tilde{R}} + \tilde{P}\delta_{\tilde{V}}] + 2 \text{tr}[\delta_{\tilde{A}}\tilde{Q}\tilde{P}] \end{aligned}$$

and

$$\begin{aligned} \Omega &\triangleq \frac{1}{2} \text{tr}[\tilde{Q} \text{cl}(\tilde{A}'^*\delta_{\tilde{P}} + \delta_{\tilde{P}}\tilde{A}') + \tilde{P} \text{cl}(\tilde{A}'\delta_{\tilde{Q}} + \delta_{\tilde{Q}}\tilde{A}'^*)] \\ &\quad + \frac{1}{2} \text{tr}[\delta_{\tilde{Q}} \text{cl}(\tilde{A}'^*\tilde{P}' + \tilde{P}'\tilde{A}') + \delta_{\tilde{P}} \text{cl}(\tilde{A}'\tilde{Q}' + \tilde{Q}'\tilde{A}'^*)] + \text{tr}[\delta_{\tilde{Q}}\tilde{R}' + \delta_{\tilde{P}}\tilde{V}']. \end{aligned}$$

Computing

$$\text{tr}[\tilde{Q}\delta_{\tilde{R}} + \tilde{P}\delta_{\tilde{V}}] = 2 \text{tr}[V_2 B_c^T P_2 \delta_{B_c}] + 2 \text{tr}[Q_2 C_c^T R_2 \delta_{C_c}] + \text{tr}[P_2 \delta_{B_c} V_2 \delta_{B_c}^T + Q_2 \delta_{C_c}^T R_2 \delta_{C_c}]$$

and

$$2 \text{tr}[\delta_{\tilde{A}}\tilde{Q}\tilde{P}] = 2 \text{tr}[Z_2 \delta_{A_c}] + 2 \text{tr}[CZ_{21}^* \delta_{B_c}] + 2 \text{tr}[Z_{12}^* B \delta_{C_c}]$$

and retaining first-order terms, we obtain (4.20).

To evaluate Ω , use (4.8) and (4.10) to replace \tilde{R}' and \tilde{V}' in the last term in Ω and write $\tilde{A}' = \tilde{A} + \delta_{\tilde{A}}$, to obtain

$$\begin{aligned} \Omega &= \frac{1}{2} \text{tr}[\tilde{Q} \text{cl}(\tilde{A}^*\delta_{\tilde{P}} + \delta_{\tilde{P}}\tilde{A}) + \tilde{P} \text{cl}(\tilde{A}\delta_{\tilde{Q}} + \delta_{\tilde{Q}}\tilde{A}^*)] \\ (4.22) \quad &\quad + \frac{1}{2} \text{tr}[\tilde{Q}(\delta_{\tilde{A}}^*\delta_{\tilde{P}} + \delta_{\tilde{P}}\delta_{\tilde{A}}) + \tilde{P}(\delta_{\tilde{A}}\delta_{\tilde{Q}} + \delta_{\tilde{Q}}\delta_{\tilde{A}}^*)] \\ &\quad - \frac{1}{2} \text{tr}[\delta_{\tilde{Q}} \text{cl}(\tilde{A}'^*\tilde{P}' + \tilde{P}'\tilde{A}') + \delta_{\tilde{P}} \text{cl}(\tilde{A}'\tilde{Q}' + \tilde{Q}'\tilde{A}'^*)]. \end{aligned}$$

Next we note that

$$(4.23) \quad \text{tr}[\tilde{Q} \text{cl}(\tilde{A}^*\delta_{\tilde{P}} + \delta_{\tilde{P}}\tilde{A})] = \text{tr}[\delta_{\tilde{P}} \text{cl}(\tilde{A}\tilde{Q} + \tilde{Q}\tilde{A}^*)].$$

To see this we observe that by arguments similar to those used in the proof of Lemma 4.4 and the fact that $\delta_{\tilde{P}}: \mathcal{D}(\tilde{A}) \rightarrow \mathcal{D}(\tilde{A}^*)$ it follows that

$$\delta_{\tilde{P}} = - \int_0^\infty e^{\tilde{A}^* t} \text{cl}(\tilde{A}^*\delta_{\tilde{P}} + \delta_{\tilde{P}}\tilde{A}) e^{\tilde{A} t} dt.$$

Now, using the technique of Lemma 4.3 with the role of \tilde{R} played by $-\text{cl}(\tilde{A}^*\delta_{\tilde{P}} + \delta_{\tilde{P}}\tilde{A})$,

we see that

$$\text{tr}[\tilde{Q} \text{cl}(\tilde{A}^* \delta_{\tilde{P}} + \delta_{\tilde{P}} \tilde{A})] = -\text{tr}[\delta_{\tilde{P}} \tilde{V}] = \text{tr}[\delta_{\tilde{P}} \text{cl}(\tilde{A} \tilde{Q} + \tilde{Q} \tilde{A}^*)].$$

Similarly, it can be shown that

$$(4.24) \quad \text{tr}[\tilde{P} \text{cl}(\tilde{A} \delta_{\tilde{Q}} + \delta_{\tilde{Q}} \tilde{A}^*)] = \text{tr}[\delta_{\tilde{Q}} \text{cl}(\tilde{A}^* \tilde{P} + \tilde{P} \tilde{A})].$$

Now substitute (4.23) and (4.24) into (4.22) and rearrange the second term in (4.22) so that

$$\begin{aligned} \Omega &= \frac{1}{2} \text{tr}[\delta_{\tilde{Q}} \text{cl}(\tilde{A}^* \tilde{P} + \tilde{P} \tilde{A}) + \delta_{\tilde{P}} \text{cl}(\tilde{A} \tilde{Q} + \tilde{Q} \tilde{A}^*)] \\ &\quad + \frac{1}{2} \text{tr}[\delta_{\tilde{Q}}(\delta_{\tilde{A}}^* \tilde{P} + \tilde{P} \delta_{\tilde{A}}) + \delta_{\tilde{P}}(\delta_{\tilde{A}} \tilde{Q} + \tilde{Q} \delta_{\tilde{A}}^*)] \\ &\quad - \frac{1}{2} \text{tr}[\delta_{\tilde{Q}} \text{cl}(\tilde{A}'^* \tilde{P}' + \tilde{P}' \tilde{A}') + \delta_{\tilde{P}} \text{cl}(\tilde{A}' \tilde{Q}' + \tilde{Q}' \tilde{A}'^*)] \\ &= -\frac{1}{2} \text{tr}[\delta_{\tilde{Q}} \text{cl}(\tilde{A}'^* \delta_{\tilde{P}} + \delta_{\tilde{P}} \tilde{A}') + \delta_{\tilde{P}} \text{cl}(\tilde{A}' \delta_{\tilde{Q}} + \delta_{\tilde{Q}} \tilde{A}'^*)]. \end{aligned}$$

Using (4.8) to obtain

$$0 = \tilde{A}' \delta_{\tilde{Q}} + \delta_{\tilde{Q}} \tilde{A}'^* + \delta_{\tilde{A}} \tilde{Q} + \tilde{Q} \delta_{\tilde{A}}^* + \delta_{\tilde{V}}$$

and (4.10) to obtain a similar relation involving \tilde{P} , we have

$$\Omega = \text{tr}[\delta_{\tilde{Q}}(\delta_{\tilde{A}}^* \tilde{P} + \tilde{P} \delta_{\tilde{A}} + \delta_{\tilde{R}})] + \text{tr}[\delta_{\tilde{P}} \tilde{Q} + \tilde{Q} \delta_{\tilde{A}}^* + \delta_{\tilde{V}}].$$

Restricting (A'_c, B'_c, C'_c) to N (see (4.14)), using Lemma 4.6 and noting that $\delta_{\tilde{A}}$ and $\delta_{\tilde{R}}$ have finite rank, it follows that there exists $c_1 > 0$ such that

$$(4.25) \quad \|\Omega\| \leq c_1 \|(\delta_{A_c}, \delta_{B_c}, \delta_{C_c})\|^2.$$

Combining Ω with the second-order terms in Λ yields the desired result. \square

LEMMA 4.8. \mathcal{A}_+ is open.

Proof. From the "generic" property of controllability and observability [62, p. 44] there exists an open neighborhood of (A_c, B_c, C_c) each of whose elements is minimal. Combining this fact with Lemma 4.5 yields the desired result. \square

LEMMA 4.9. Q_2 and P_2 are positive definite.

Proof. First note that expanding the $\mathbf{R}^{n \times n}$ -component of the Lyapunov equation (4.8) yields (4.50) below. By a minor extension of results from [66] or [67], (4.50) can be rewritten as

$$0 = (A_c + B_c C Q_{12} Q_2^+) Q_2 + Q_2 (A_c + B_c C Q_{12} Q_2^+)^T + B_c V_2 B_c^T,$$

where Q_2^+ is the Moore-Penrose or Drazin generalized inverse of Q_2 . Next note that since (A_c, B_c) is controllable then so is $(A_c + B_c C Q_{12} Q_2^+, B_c V_2^{1/2})$. Now, since Q_2 and $B_c V_2 B_c^T$ are nonnegative definite, it follows from [62, Lemma 12.2] that Q_2 is positive definite. Similar arguments show that P_2 is positive definite. \square

Having established Lemmas 4.1-4.9, we can now proceed with the proof of the Main Theorem. Let $(A_c, B_c, C_c) \in \mathcal{A}_+$ be as in the Main Theorem and consider (4.19) with (A'_c, B'_c, C'_c) confined to \mathcal{A}_+ . Because $\mathcal{L}: \mathbf{R}^{n \times n} \times \mathbf{R}^{n \times 1} \times \mathbf{R}^{m \times n} \rightarrow \mathbf{R}$ is a bounded linear functional and \mathcal{A}_+ is open, the convergence in (4.21) implies that \mathcal{L} is precisely the Frechet derivative of J with respect to (A_c, B_c, C_c) . Since \mathcal{A}_+ is open, the optimality of (A_c, B_c, C_c) implies

$$(4.26) \quad \mathcal{L}(\delta_{A_c}, \delta_{B_c}, \delta_{C_c}) = 0$$

for all $(\delta_{A_c}, \delta_{B_c}, \delta_{C_c})$. Clearly, (4.26) is equivalent to

$$(4.27) \quad Z_2 = 0,$$

$$(4.28) \quad V_2 B_c^T P_2 + C Z_{21}^* = 0,$$

$$(4.29) \quad Q_2 C_c^T R_2 + Z_{12}^* B = 0.$$

Thus, B_c and C_c are given by

$$(4.30) \quad B_c = -P_2^{-1} Z_{21} C^* V_2^{-1},$$

$$(4.31) \quad C_c = -R_2^{-1} B^* Z_{12} Q_2^{-1}.$$

Although B_c and C_c are now determined in terms of \tilde{Q} and \tilde{P} , A_c remains to be found. Moreover, \tilde{Q} and \tilde{P} themselves depend (via (4.8) and (4.10)) on B_c and C_c . Hence our task now is to consolidate and simplify (4.7)-(4.10), (4.27), (4.30) and (4.31) to obtain the more tractable conditions (3.9)-(3.18). To this end let us define new variables

$$(4.32a, b) \quad Q \triangleq Q_1 - Q_{12} Q_2^{-1} Q_{12}^*, \quad P \triangleq P_1 - P_{12} P_2^{-1} P_{12}^*,$$

$$(4.33a, b) \quad \hat{Q} \triangleq Q_{12} Q_2^{-1} Q_{12}^*, \quad \hat{P} \triangleq P_{12} P_2^{-1} P_{12}^*.$$

Clearly, \hat{Q} and \hat{P} are nonnegative definite and have finite rank. Since by Lemma 4.2 $\tilde{Q}, \tilde{P} \in \mathcal{B}_1(\mathcal{H})$, it can be seen that $Q, P \in \mathcal{B}_1(\mathcal{H})$, which implies $Q, P \in \mathcal{B}(\mathcal{H})$. To show that Q and P are nonnegative definite, note that Q is the $\mathcal{B}(\mathcal{H})$ -component of the nonnegative-definite operator $\mathcal{Q} \tilde{Q} \mathcal{Q}^* \in \mathcal{B}(\mathcal{H})$, where

$$\mathcal{Q} \triangleq \begin{bmatrix} I_{\mathcal{H}} & -Q_{12} Q_2^{-1} \\ 0 & -I_{n_c} \end{bmatrix}.$$

Similarly, P is nonnegative definite.

From the domain conditions (4.7) and (4.9) it follows that

$$(4.34a, b) \quad Q_1: \mathcal{D}(A^*) \rightarrow \mathcal{D}(A), \quad P_1: \mathcal{D}(A) \rightarrow \mathcal{D}(A^*),$$

$$(4.35a, b) \quad Q_{12}: \mathcal{R}^{n_c} \rightarrow \mathcal{D}(A), \quad P_{12}: \mathcal{R}^{n_c} \rightarrow \mathcal{D}(A^*),$$

which lead to (3.12) and (3.13).

Next note that (4.27) is equivalent to (3.8) with

$$(4.36a, b) \quad G \triangleq Q_2^{-1} Q_{12}^*, \quad \Gamma \triangleq -P_2^{-1} P_{12}^*$$

and that (3.7) holds with

$$(4.37) \quad M \triangleq Q_2 P_2.$$

Since Q_2 and P_2 are positive definite, Lemma 2.6 implies that M is positive semisimple. We can also define $\tau = G^* \Gamma$ which, by (3.8) satisfies $\tau^2 = \tau$. It is helpful to note the identities

$$(4.38a, b) \quad \hat{Q} = Q_{12} G = G^* Q_{12}^*, \quad \hat{P} = -P_{12} \Gamma = -\Gamma^* P_{12}^*,$$

$$(4.39a, b) \quad \hat{Q} = G^* Q_2 G, \quad \hat{P} = \Gamma^* P_2 \Gamma,$$

$$(4.40a, b) \quad \tau G^* = G^*, \quad \Gamma \tau = \Gamma,$$

$$(4.41a, b) \quad \hat{Q} = \tau \hat{Q}, \quad \hat{P} = \hat{P} \tau,$$

$$(4.42) \quad \hat{Q} \hat{P} = -Q_{12} P_{12}^*.$$

From (3.8) and (2.1) it follows that

$$(4.43a, b) \quad \rho(G) = \rho(\Gamma) = n_c,$$

$$(4.44a, b) \quad \rho(Q_{12}) = \rho(P_{12}) = n_c.$$

Hence, (2.2) and (4.38) imply $n_c = \rho(Q_{12}) + \rho(G) - n_c \leq \rho(\hat{Q}) \leq \rho(Q_{12}) = n_c$ which yields (3.14a). Similarly, (3.14b) holds and (3.14c) follows from (2.2) and (4.42).

Using (4.38) and (4.39), the components of \hat{Q} and \hat{P} can be written in terms of G, Γ, Q, P, \hat{Q} and \hat{P} as

$$(4.45) \quad Q_1 = Q + \hat{Q}, \quad P_1 = P + \hat{P},$$

$$(4.46) \quad Q_{12} = \hat{Q}\Gamma^*, \quad P_{12} = -\hat{P}G^*,$$

$$(4.47) \quad Q_2 = \Gamma\hat{Q}\Gamma^*, \quad P_2 = G\hat{P}G^*.$$

Now (3.10) and (3.11) can be obtained by substituting (4.45)–(4.47) into (4.30) and (4.31).

Expanding the $\mathcal{B}(\mathcal{H})$, $\mathcal{B}(\mathbb{R}^{n_c}, \mathcal{H})$ and $\mathbb{R}^{n_c \times n_c}$ components of (4.8) and (4.10) yields

$$(4.48) \quad 0 = AQ_1 + Q_1A^* + BC_cQ_{12}^* + Q_{12}(BC_c)^* + V_1,$$

$$(4.49) \quad 0 = AQ_{12} + Q_{12}A_c^T + BC_cQ_2 + Q_1(B_cC)^*,$$

$$(4.50) \quad 0 = A_cQ_2 + Q_2A_c^T + B_cCQ_{12} + Q_{12}^*(B_cC)^* + B_cV_2B_c^T,$$

$$(4.51) \quad 0 = A^*P_1 + P_1A + (B_cC)^*P_{12}^* + P_{12}B_cC + R_1,$$

$$(4.52) \quad 0 = P_{12}A_c + A^*P_{12} + (B_cC)^*P_2 + P_1BC_c,$$

$$(4.53) \quad 0 = A_c^TP_2 + P_2A_c + (BC_c)^*P_{12} + P_{12}^*BC_c + C_c^TR_2C_c.$$

Substituting (4.45)–(4.47) into (4.48)–(4.53), using the identities

$$B_cC = \Gamma Q\bar{\Sigma}, \quad BC_c = -\Sigma PG^*,$$

$$B_cV_2B_c^T = \Gamma Q\bar{\Sigma}Q\Gamma^*, \quad C_c^TR_2C_c = GP\Sigma PG^*,$$

and defining

$$A_Q \triangleq A - Q\bar{\Sigma}, \quad A_P \triangleq A - \Sigma P,$$

we obtain

$$(4.54) \quad 0 = AQ + QA^* + A_P\hat{Q} + \hat{Q}A_P + V_1,$$

$$(4.55) \quad 0 = [A_P\hat{Q} + Q\bar{\Sigma}Q + \hat{Q}(\Gamma^*A_c^TG + \bar{\Sigma}Q)]\Gamma^*,$$

$$(4.56) \quad 0 = \Gamma[G^*A_c\Gamma\hat{Q} + Q\bar{\Sigma}\hat{Q} + Q\bar{\Sigma}Q + \hat{Q}(\Gamma^*A_c^TG + \bar{\Sigma}Q)]\Gamma^*,$$

$$(4.57) \quad 0 = A^*P + PA + A_Q^*\hat{P} + \hat{P}A_Q + R_1,$$

$$(4.58) \quad 0 = -[A_Q^*\hat{P} + P\Sigma P + \hat{P}(G^*A_c\Gamma + \Sigma P)]G^*,$$

$$(4.59) \quad 0 = G[\Gamma^*A_c^TG\hat{P} + P\Sigma\hat{P} + P\Sigma P + \hat{P}(G^*A_c\Gamma + \Sigma P)]G^*.$$

We are now in a position to determine A_c by computing (4.56) – $\Gamma(4.55)$ which yields (3.9). Alternatively, A_c can be obtained by computing (4.59) + $G(4.58)$. As mentioned in § 3, (3.9) is valid since $G^*: \mathbb{R}^{n_c} \rightarrow \mathcal{D}(A)$ and A_c^T is given by (3.26).

Next we substitute the expressions for A_c and A_c^T into (4.55), (4.56), (4.58) and (4.59) and compute the relations (4.55) G , $G^*(4.56)G$, $-(4.58)\Gamma$ and $\Gamma^*(4.59)\Gamma$ to obtain, respectively,

$$(4.60) \quad 0 = [A_P\hat{Q} + \hat{Q}A_P^* + Q\bar{\Sigma}Q]\tau^*,$$

$$(4.61) \quad 0 = \tau[A_P\hat{Q} + \hat{Q}A_P^* + Q\bar{\Sigma}Q]\tau^*,$$

$$(4.62) \quad 0 = [A_Q^*\hat{P} + \hat{P}A_Q + P\Sigma P]\tau,$$

$$(4.63) \quad 0 = \tau^*[A_Q^*\hat{P} + \hat{P}A_Q + P\Sigma P]\tau.$$

Note that (4.60)–(4.63) are *equivalent* to (4.55), (4.56), (4.58) and (4.59) since G and Γ have full rank. Since $(4.61) = \tau(4.60)$ and $(4.63) = \tau^*(4.62)$, (4.61) and (4.63) are superfluous and can be omitted. Thus we have derived (3.17) and (3.18).

To obtain (3.15) and (3.16) we need only compute the relations $(4.54) + \tau(4.60) - (4.60) - (4.60)^*$ and $(4.57) + \tau^*(4.62) - (4.62) - (4.62)^*$ and use (4.41).

Finally, to show that the preceding development entails no loss of generality in the optimality conditions we now use (3.9)–(3.18) to obtain (4.7)–(4.10) and (4.27)–(4.29). Let $A_\infty, B_\infty, C_\infty, G, \Gamma, \tau, Q, P, \hat{Q}, \hat{P}$ be as in the theorem statement and define $Q_1, Q_{12}, Q_2, P_1, P_{12}, P_2$ by (4.45)–(4.47). Note that (3.12) and (3.13) imply (4.34) and (4.35) and hence (4.7) and (4.9). Using (3.8), (3.10), (3.11) and (3.22) it is easy to verify (4.27)–(4.29). Finally, substitute (4.32), (4.33) and (4.36) into (3.15)–(3.18), reverse the steps taken earlier in the proof and use (3.9)–(3.11) to obtain (4.8) and (4.10), which completes the proof. \square

5. Concluding remarks. This paper has considered the problem of quadratically optimal, steady-state, fixed-order dynamic compensation for linear infinite-dimensional systems. The Main Theorem presents the stationarity conditions of the optimization problem in a highly simplified and rigorous form. The “optimal projection equations” (3.15)–(3.18) (or, equivalently, (3.27)–(3.30)) of the Main Theorem reveal the essential structure of the first-order necessary conditions and display the central role played by the optimal projection τ . The relationship of the Main Theorem to the standard finite-dimensional steady-state *LQG* problem can be demonstrated by replacing τ with the identity matrix and noting that (3.27) and (3.28) reduce immediately to the familiar pair of operator Riccati equations and that (3.29) and (3.30) yield the controllability and observability gramians of the controller.

Inasmuch as the Main Theorem is a fundamental generalization of classical steady-state *LQG* theory, a number of issues must be reexamined. Hence, in conclusion we should like to point out some possible extensions of the Main Theorem along with directions for further research.

1. *Sufficiency theory.* Although sufficient conditions for the existence of an optimal compensator were not investigated in this paper, auxiliary conditions based upon the structure of (3.15)–(3.18) could perhaps be imposed upon Q, P, \hat{Q} and \hat{P} to single out the global optimum from amongst the local minima. This would be similar to the situation in *LQG* theory where, under stabilizability and detectability hypotheses, optimal stabilizing Q and P are identified as the unique nonnegative-definite solutions of the pair of algebraic Riccati equations.

2. *Stabilizability.* Just as in the full-order *LQG* problem, one would expect a natural relationship between the structure of the optimal solution and stabilizability/detectability hypotheses. The results of [41], [42] and [68] could serve as a starting point in this regard.

3. *Numerical algorithms.* In practical situations, the distributed parameter system would be replaced by a high-order discretized model for which the matrix version (rather than the operator version) of the optimal projection equations could be solved numerically. A numerical algorithm for solving the matrix version of the optimal projection equations has been developed in [32] and [34]. The proposed computational scheme is fundamentally quite different from gradient search algorithms [17], [18], [21], [22], [24], [25], [28], [30] in that it operates through direct solution of the optimal projection equations by iterative refinement of the optimal projection.

4. *Convergence.* One of the principal uses for the optimal projection equations will be to understand the relationship between fixed-order dynamic-compensator

designs which are optimal with respect to approximate models and the optimal fixed-order dynamic compensator for the distributed parameter system itself. By considering a sequence of n th-order approximate models which converge to the distributed parameter system, conditions would be sought guaranteeing that the sequence of fixed-order compensators based on each approximate model approach the optimal dynamic compensator based upon the distributed parameter system (see [38]–[40]). This approach is analogous to the convergence results obtained in [7], [8] with the major difference being that the optimal projection equations permit the order of the compensator to remain fixed in accordance with real-world implementation constraints whereas in [7]–[9] the order of the compensator increases without bound.

5. *Unbounded control and observation.* An important generalization of the problem considered in this paper involves the case in which the input and output operators B and C are unbounded. The mathematical details for this problem are considerably more complex (see, e.g., [69]).

6. *Singular observation noise/singular control weighting.* As pointed out in [22], [33], [36] the assumptions of nonsingular control weighting and nonsingular observation noise preclude the use of direct output feedback as in

$$(5.1) \quad u(t) = C_c x_c(t) + D_c y(t)$$

since J is undefined unless

$$\text{tr} [D_c^T R_2 D_c V_2] = 0 (\Leftrightarrow R_2 D_c V_2 = 0).$$

Although with due attention to (5.1) direct output feedback can be used in the singular case, the nature of the problem forebodes all of the difficulties associated with the singular *LQG* problem. Note that the deterministic output feedback problem [70], when viewed in this context, is highly singular.

7. *Discrete-time system/discrete-time compensator.* Digital implementation can be modelled by a discrete-time compensator with control of a continuous-time system facilitated by sampling and reconstruction devices. See [71], [73] for results in this direction.

8. *Cross weighting/correlated disturbance and observation noise.* This extension is straightforward and entirely analogous to the *LQG* case (see, e.g., [18, p. 351]).

Acknowledgments. We wish to thank Ardeth P. Grant for excellent and careful typing of the original manuscript version of this paper. Word processing and revision support was provided by Harris Corporation, GASD.

REFERENCES

- [1] M. J. BALAS, *Toward a more practical control theory for distributed parameter systems*, in Control and Dynamic Systems, C. T. Leondes, ed., Advances in Theory and Applications, 19, Academic Press, New York, 1982.
- [2] M. ATHANS, *Toward a practical theory of distributed parameter systems*, IEEE Trans. Automat. Control., AC-15 (1970), pp. 245–247.
- [3] S. A. REIBLE, *Acoustoelectric convolver technology for spread-spectrum communications*, IEEE Trans. Microwave Theory Tech., MTT-29 (1981), pp. 463–473.
- [4] R. F. CURTAIN AND A. J. PRITCHARD, *Infinite Dimensional Linear Systems Theory*, Springer-Verlag, New York, 1978.
- [5] J. S. GIBSON, *The Riccati integral equations for optimal control problems on Hilbert spaces*, this Journal, 17 (1979), pp. 537–565.

- [6] A. V. BALAKRISHNAN, *Applied Functional Analysis*, Springer-Verlag, New York, 1981.
- [7] J. S. GIBSON, *An analysis of optimal modal regulation: convergence and stability*, this Journal, 19 (1981), pp. 686-707.
- [8] ———, *Linear-quadratic optimal control of hereditary differential systems: infinite dimensional Riccati equations and numerical approximations*, this Journal, 21 (1983), pp. 95-139.
- [8a] H. T. BANKS AND K. KUNISCH, *The linear regulator problem for parabolic systems*, this Journal, 22 (1984), pp. 684-698.
- [9] H. T. BANKS, I. G. ROSEN AND K. ITO, *A spline based technique for computing Riccati operators and feedback controls in regulator problems for delay equations*, SIAM J. Sci. Stat. Comput., 5 (1984), pp. 830-855.
- [10] M. AOKI, *Control of large-scale dynamic systems by aggregation*, IEEE Trans. Automat. Control, AC-13 (1968), pp. 246-253.
- [11] D. A. WILSON, *Optimum solution of model-reduction problems*, Proc. IEE, 117 (1970), pp. 1161-1165.
- [12] B. C. MOORE, *Principal component analysis in linear systems: controllability, observability, and model reduction*, IEEE Trans. Automat. Control, AC-26 (1981), pp. 17-32.
- [13] R. E. SKELTON AND A. YOUSUFF, *Component cost analysis of large scale systems*, in Control and Dynamic Systems, C. T. Leondes, ed., Academic Press, New York, 1982.
- [14] E. A. JONCKHEERE AND L. M. SILVERMAN, *A new set of invariants for linear systems—application to reduced-order compensator design*, IEEE Trans. Automat. Control, AC-28 (1983), pp. 953-964.
- [15] A. YOUSUFF AND R. E. SKELTON, *Controller reduction by component cost analysis*, IEEE Trans. Automat. Control, AC-29 (1984), pp. 520-530.
- [16] T. L. JOHNSON AND M. ATHANS, *On the design of optimal constrained dynamic compensators for linear constant systems*, IEEE Trans. Automat. Control, AC-15 (1970), pp. 658-660.
- [17] W. S. LEVINE, T. L. JOHNSON AND M. ATHANS, *Optimal limited state variable feedback controllers for linear systems*, IEEE Trans. Automat. Control, AC-16 (1971), pp. 785-793.
- [18] K. KWAKERNAAK AND R. SIVAN, *Linear Optimal Control Systems*, Wiley-Interscience, New York, 1972.
- [19] D. B. ROM AND P. E. SARACHIK, *The design of optimal compensators for linear constant systems with inaccessible states*, IEEE Trans. Automat. Control, AC-18 (1973), pp. 509-512.
- [20] M. SIDAR AND B.-Z. KURTARAN, *Optimal low-order controllers for linear stochastic systems*, Int. J. Control, 22 (1975), pp. 377-387.
- [21] J. M. MENDEL AND J. FEATHER, *On the design of optimal time-invariant compensators for linear stochastic time-invariant systems*, IEEE Trans. Automat. Control, AC-20 (1975), pp. 653-657.
- [22] S. BASUTHAKUR AND C. H. KNAPP, *Optimal constant controllers for stochastic linear systems*, IEEE Trans. Automat. Control, AC-20 (1975), pp. 664-666.
- [23] R. B. ASHER AND J. C. DURRETT, *Linear discrete stochastic control with a reduced-order dynamic compensator*, IEEE Trans. Automat. Control, AC-21 (1976), pp. 626-627.
- [24] W. J. NAEIJE AND O. H. BOSGRA, *The design of dynamic compensators for linear multivariable systems*, 1977 IFAC, Fredericton, New Brunswick, Canada, pp. 205-212.
- [25] H. R. SIRISENA AND S. S. CHOI, *Design of optimal constrained dynamic compensators for non-stationary linear stochastic systems*, Int. J. Contr., 25 (1977), pp. 513-524.
- [26] P. J. BLANVILLAIN AND T. L. JOHNSON, *Specific-optimal control with a dual minimal-order observer-based compensator*, Int. J. Contr., 28 (1978), pp. 277-294.
- [27] ———, *Invariants of optimal minimal-order observer-based compensators*, IEEE Trans. Automat. Control, AC-23 (1978), pp. 473-474.
- [28] C. J. WENK AND C. H. KNAPP, *Parameter optimization in linear systems with arbitrarily constrained controller structure*, IEEE Trans. Automat. Control, AC-25 (1980), pp. 496-500.
- [29] J. O'REILLY, *Optimal low-order feedback controllers for linear discrete-time systems*, in Control and Dynamic Systems 16, C. T. Leondes, ed., Academic Press, New York, 1980.
- [30] D. P. LOOZE AND N. R. SANDELL, JR., *Gradient calculations for linear quadratic fixed control structure problems*, IEEE Trans. Automat. Control, AC-25 (1980), pp. 285-288.
- [31] D. C. HYLAND, *Optimality conditions for fixed-order dynamic compensation of flexible spacecraft with uncertain parameters*, AIAA 20th Aerospace Sciences Mtg., paper 82-0312, Orlando, FL, Jan. 1982.
- [32] ———, *The optimal projection approach to fixed-order compensation: Numerical methods and illustrative results*, AIAA 21st Aerospace Sciences Mtg., paper 83-0303, Reno, NV, Jan. 1983.
- [33] D. C. HYLAND AND D. S. BERNSTEIN, *Explicit optimality conditions for fixed-order dynamic compensation*, Proc. IEEE Conference on Decision and Control, San Antonio, TX, Dec. 1983, pp. 161-165.
- [34] D. C. HYLAND, *Comparison of various controller-reduction methods: Suboptimal versus optimal projection*, Proc. AIAA Dynamics Specialists Conference, Palm Springs, CA, May 1984, pp. 382-389.

- [35] D. S. BERNSTEIN AND D. C. HYLAND, *The optimal projection equations for fixed-order dynamic compensation of distributed parameter systems*, Proc. AIAA Dynamics Specialists Conference, Palm Springs, CA, May 1984, pp. 396-400.
- [36] ———, *The optimal projection equations for fixed-order dynamic compensation*, IEEE Trans. Automat. Control, AC-29 (1984), pp. 1034-1037.
- [37] ———, *The optimal projection approach to model reduction and the relationship between the methods of Wilson and Moore*, Proc. IEEE Conference on Decision and Control, Las Vegas, NV, Dec. 1984, pp. 120-126.
- [37a] ———, *The optimal projection equations for model reduction and the relationships among the methods of Wilson, Skelton and Moore*, to appear.
- [38] T. L. JOHNSON, *Optimization of low order compensators for infinite dimensional systems*, Proc. 9th IFIP Symposium on Optimization Techniques, Warsaw, Poland, September 1979, pp. 394-401.
- [39] R. K. PEARSON, *Optimal fixed-form compensators for large space structures*, in ACOSS SIX (Active Control of Space Structures), RADC-TR-81-289, Final Technical Report, Rome Air Development Center, Griffiss AFB, New York, 1981.
- [40] ———, *Optimal velocity feedback control of flexible structures*, Ph.D. dissertation, Dept. Electrical Engineering and Computer Science, Massachusetts Institute of Technology, Cambridge, 1982.
- [41] R. F. CURTAIN, *Compensators for infinite-dimensional linear systems*, J. Franklin Inst., 315 (1983), pp. 331-346.
- [42] J. M. SCHUMACHER, *A direct approach to compensator design for distributed parameter systems*, this Journal, 21 (1983), pp. 823-836.
- [43] D. L. RUSSELL, *Linear stabilization of the linear oscillator in Hilbert space*, J. Math. Anal. Appl., 25 (1969), pp. 663-675.
- [44] ———, *Decay rates for weakly damped systems in Hilbert space obtained with control theoretic methods*, J. Differential Equations, 19 (1975), pp. 344-370.
- [45] M. J. BALAS, *Modal control of certain flexible dynamic systems*, this Journal, 16 (1978), pp. 450-462.
- [46] ———, *Feedback control of flexible systems*, IEEE Trans. Automat. Control, AC-23 (1978), pp. 673-679.
- [47] J. S. GIBSON, *A note on stabilization of infinite dimensional linear oscillators by compact feedback*, this Journal, 18 (1980), pp. 311-316.
- [48] M. J. BALAS, *Trends in large space structure control theory: Fondest hopes, wildest dreams*, IEEE Trans. Automat. Control, AC-24 (1982), pp. 522-535.
- [49] T. L. JOHNSON, *Progress in modelling and control of flexible spacecraft*, J. Franklin Inst., 315 (1983), pp. 495-520.
- [49a] M. J. BALAS, *Feedback control of dissipative hyperbolic distributed parameter systems with finite dimensional controllers*, J. Math. Anal. Appl., 98 (1984), pp. 1-24.
- [50] T. KATO, *Perturbation Theory for Linear Operators*, Springer-Verlag, New York, 1966.
- [51] J. R. RINGROSE, *Compact Non-Self-Adjoint Operators*, Van Nostrand Reinhold, London, 1971.
- [52] I. C. GOHBERG AND M. G. KREIN, *Introduction to the Theory of Linear Nonselfadjoint Operators*, Translations of Mathematical Monographs, Vol. 18, American Mathematical Society, Providence, RI, 1966.
- [52a] M. S. BRODSKII, *Triangular and Jordan Representations of Linear Operators*, Translations of Mathematical Monographs, Vol. 32, American Mathematical Society, Providence, RI, 1971.
- [53] I. GOHBERG AND S. GOLDBERG, *Basic Operator Theory*, Birkhauser, Boston, 1981.
- [54] F. R. GANTMACHER, *The Theory of Matrices*, Vol. 1, Chelsea, New York, 1977.
- [55] C. R. RAO AND S. K. MITRA, *Generalized Inverse of Matrices and Its Applications*, John Wiley, New York, 1971.
- [56] B. NOBLE AND J. W. DANIEL, *Applied Linear Algebra*, Second edition, Prentice-Hall, Englewood Cliffs, NJ, 1977.
- [57] R. F. CURTAIN AND A. J. PRITCHARD, *Functional Analysis in Modern Applied Mathematics*, Academic Press, London, 1977.
- [58] S. CHAKRABARTI, B. B. BATTACHARYYA AND M. N. S. SWAMY, *On simultaneous diagonalization of a collection of hermitian matrices*, Matrix and Tensor Quart., 29 (1978), pp. 35-54.
- [59] C. T. MULLIS AND R. A. ROBERTS, *Synthesis of minimum roundoff noise fixed point digital filters*, IEEE Trans. Circ. Syst., CAS-23 (1976), pp. 551-562.
- [60] A. J. LAUB, *Computation of balancing transformation*, Proc. 1980 Joint Automation Control Conference, San Francisco, CA, Aug. 1980.
- [61] E. JONCKHEERE, *Open-loop and closed loop approximations of linear systems and associated balanced realizations*, 1982 Symposium on Circuits and Systems, Rome, May 1982.
- [62] W. M. WONHAM, *Linear Multivariable Control: A Geometric Approach*, Springer-Verlag, New York, 1974.

- [63] D. C. LAY, *Spectral properties of generalized inverses of linear operators*, SIAM J. Appl. Math., 29 (1975), pp. 103-109.
- [64] S. L. CAMPBELL AND C. D. MEYER, JR., *Generalized Inverses of Linear Transformations*, Pitman, London, 1979.
- [65] P. ROBERT, *On the group-inverse of a linear transformation*, J. Math. Anal. Appl., 22 (1968), pp. 658-669.
- [66] A. ALBERT, *Conditions for positive and nonnegative definiteness in terms of pseudo inverse*, SIAM J. Appl. Math., 17 (1969), pp. 434-440.
- [67] E. KREINDLER AND A. JAMESON, *Conditions for nonnegativeness of partitioned matrices*, IEEE Trans. Automat. Control, AC-17 (1972), pp. 147-148.
- [68] C. N. NETT, C. A. JACOBSON AND M. J. BALAS, *Fractional representation theory: Robustness results with applications to finite dimensional control of a class of linear distributed systems*, IEEE Conference on Decision and Control, San Antonio, TX, Dec. 1983.
- [69] R. F. CURTAIN, *Finite-dimensional compensators for parabolic distributed systems with unbounded control and observation*, this Journal, 22 (1984), pp. 255-276.
- [70] W. S. LEVINE AND M. ATHANS, *On the determination of the optimal constant output feedback gains for linear multivariable systems*, IEEE Trans. Automat. Control, AC-15 (1970), pp. 44-48.
- [71] M. J. BALAS, *The structure of discrete-time finite-dimensional control of distributed parameter systems*, Proc. IEEE International Large Scale Systems Symposium, Virginia Beach, VA, 1982.
- [72] M. S. GHAVSI AND J. J. KELLY, *Introduction to Distributed-Parameter Networks with Application to Integrated Circuits*, Holt, Rinehart and Winston, New York, 1968.
- [73] D. S. BERNSTEIN, L. D. DAVIS AND D. C. HYLAND, *The optimal projection equations for reduced-order, discrete-time modelling, estimation and control*, submitted for publication.

April 1988

**Finite-Dimensional Approximation for
Optimal Fixed-Order Compensation of
Distributed Parameter Systems**

by

Dennis S. Bernstein*
Government Aerospace Systems Division
Harris Corporation
MS 22/4848
Melbourne, FL 32902

I. Gary Rosen**
Department of Mathematics
University of Southern California
Los Angeles, CA 90089

Abstract

In controlling distributed parameter systems it is often desirable to obtain low-order, finite-dimensional controllers in order to minimize real-time computational requirements. Standard approaches to this problem employ model/controller reduction techniques in conjunction with LQG theory. In this paper we consider the finite-dimensional approximation of the infinite-dimensional Bernstein/Hyland optimal projection theory. Our approach yields fixed-finite-order controllers which are optimal with respect to high-order, approximating, finite-dimensional plant models. We illustrate the technique by computing a sequence of first-order controllers for one-dimensional, single-input/single-output, parabolic (heat/diffusion) and hereditary systems using spline-based, Ritz- Galerkin, finite element approximation. Our numerical studies indicate convergence of the feedback gains with less than 2% performance degradation over full-order LQG controllers for the parabolic system and 10% degradation for the hereditary system.

* This research supported in part by the Air Force Office of Scientific Research under contract F49620-86-C-0002

** This research supported in part by the Air Force Office of Scientific Research under grant AFOSR-87-0351. Part of this research was carried out while this author was a visiting scientist at the Institute for Computer Applications in Science and Engineering (ICASE) at the NASA Langley Research Center in Hampton, VA which is operated under NASA contracts NAS1-17070 and NAS1-18107.

1. Introduction

Approximation methods for the optimal control of distributed parameter systems have been widely studied. In particular, the approach taken in [1-12] involves approximating the original distributed parameter system by a sequence of finite-dimensional systems and then using finite-dimensional control-design techniques to obtain a sequence of approximating, sub-optimal control laws, observers, or compensators. Furthermore, in these treatments it was demonstrated that if the open-loop system is approximated appropriately, then it is possible to guarantee convergence of the sequence of sub-optimal controllers, observers, or compensators, respectively, to the optimal controller, observer, or compensator for the original infinite-dimensional system. In addition, it can also be shown that when the approximating sub-optimal control laws or estimators are applied to the original system, near-optimal performance can frequently be obtained. These ideas were pursued in the context of both open- and closed-loop control, in both continuous and discrete-time, and for both full-state-feedback control and LQG (i.e., Kalman-filter-based) state estimation and compensation.

In practical situations, however, it is often of interest to obtain the simplest (i.e., the lowest order) controller which provides a given, desired feedback performance. This is usually achieved in one of two ways. Either the plant approximation order is reduced prior to controller design, or, alternatively, reduction techniques are applied to a given high-order control law. Unfortunately, the former approach may result in undesirable spillover effects while the latter may yield low-order controllers of low authority which perform unacceptably. In fact, with the second approach, this may occur even when a suitable controller is known to exist. For example, as is shown in [13], controller reduction techniques may even destabilize the closed-loop system.

A third, more direct approach involves fixing the controller order a priori, and then optimizing a performance criterion over the class of fixed-order controllers. In a finite-dimensional setting, a set of necessary conditions in the form of four coupled matrix equations (as a direct extension of the pair of the separated Riccati equations of LQG theory) which characterize the optimal fixed-order compensator was derived in [14]. These four equations are coupled via an oblique projection (idempotent) matrix. In the full-order case, this projection becomes the identity thus effectively eliminating the additional two equations, and the necessary conditions reduce to the standard LQG Riccati equations.

The notion that this direct (i.e., fixed-finite-order) approach can be applied to distributed parameter systems was first suggested by Johnson in [15] and further developed in [16] and [17]. To realize such an approach, however, would require a suitable generalization of the optimality conditions for the finite-dimensional fixed-order theory. This result was subsequently obtained in [18] where the matrix optimal projection equations obtained in [14] for finite-dimensional systems were extended to a set of four coupled *operator* Riccati and Lyapunov equations characterizing optimal fixed-finite-order controllers for infinite-dimensional systems.

In developing numerical schemes to actually compute fixed-finite-order compensators for infinite-dimensional systems, one might consider an approach wherein LQG reduction procedures are applied to a sequence of controllers obtained by using finite-dimensional full-order design techniques in conjunction with high-order finite-dimensional plant approximations. However, such an approach is unappealing for two reasons. First, since such methods are not predicated on the minimization of a performance index, prospects for convergence are slim. And, second, controller-reduction methods have not proven to be reliable in producing stabilizing compensators (see, for example, [13]).

Hence, on the other hand, we develop an abstract approximation framework (and ultimately computational schemes) which combine the infinite-dimensional optimal projection theory of [18] with the approximation ideas developed in [9-12] for infinite-dimensional LQG problems. More precisely, our approach involves constructing a sequence of approximating finite-dimensional subspaces of the original, underlying, infinite-dimensional Hilbert state space along with corresponding sequences of bounded linear operators which approximate the given input, output, and system operators. Then, by choosing bases for these approximating subspaces and applying the finite-dimensional optimal projection theory, a sequence of matrix equations characterizing a sequence of approximating optimal, fixed-finite-order compensators for the distributed system is obtained. Finally, numerical techniques for solving the matrix optimal projection equations (for example, the homotopic continuation algorithm described in [19] and [20]) can be used to compute the sequence of approximating gains.

Our primary aim in this paper is to describe the general approach we are proposing, to discuss its implementation, and to demonstrate its feasibility and practicality. We offer no convergence arguments here, but rather reserve them for a more theoretical paper to follow. Instead, we consider the application of our technique to two examples. One involves the control of a one-dimensional,

single-input, single-output parabolic (heat/diffusion) system while the other involves a single-input single-output one-dimensional hereditary control system. These relatively simple examples have been used throughout the distributed parameter control literature to illustrate the application of new theories and techniques. A detailed discussion of the application of our ideas to more complex control systems, for instance, the vibration control of flexible structures, will also appear elsewhere. We use spline-based Ritz-Galerkin finite element schemes to approximate the open-loop systems (one for which convergence can be demonstrated in the LQG case) and present and discuss some of the numerical results which we have obtained using our general approximation framework.

We now outline the remainder of the paper. In Section 2 we briefly review the infinite-dimensional optimal projection theory from [18], describe the approximation framework, and derive the corresponding equivalent matrix equations and feedback gains which characterize the approximating fixed-finite-order compensator. In Section 3 we consider the examples, construct the approximation schemes, and discuss our numerical findings. Section 4 contains a summary and some concluding remarks.

2. Optimal Projection Theory and Finite-Dimensional Approximation

We consider the following fixed-finite-order dynamic-compensation problem. Given the infinite-dimensional control system

$$\dot{x}(t) = Ax(t) + Bu(t) + H_1 w(t) \quad (2.1)$$

with measurements

$$y(t) = Cx(t) + H_2 w(t), \quad (2.2)$$

where $t \in [0, \infty)$, design a finite-dimensional, n_c th-order dynamic compensator

$$\dot{x}_c(t) = A_c x_c(t) + B_c y(t), \quad (2.3)$$

$$u(t) = C_c x_c(t) \quad (2.4)$$

which minimizes the steady-state performance criterion

$$J(A_c, B_c, C_c) \triangleq \lim_{t \rightarrow \infty} \mathbb{E}[\langle R_1 x(t), x(t) \rangle + u(t)^T R_2 u(t)]. \quad (2.5)$$

For convenience we denote the infinite-dimensional plant by Π ; that is,

$$\Pi \triangleq \{A, B, C, R_1, R_2, V_1, V_2\}.$$

Here $x(t)$ lies in a real, separable Hilbert space \mathcal{X} with inner product $\langle \cdot, \cdot \rangle$, $A : \text{Dom}(A) \subset \mathcal{X} \rightarrow \mathcal{X}$ is a closed, densely defined operator which generates a C_0 semigroup $\{T(t) : t \geq 0\}$ of bounded linear operators on \mathcal{X} , $B \in \mathcal{L}(\mathbb{R}^m, \mathcal{X})$, and $C \in \mathcal{L}(\mathcal{X}, \mathbb{R}^l)$. We assume that the state and measurement are corrupted by a white noise signal $w(t)$ in a real, separable Hilbert space $\hat{\mathcal{X}}$ (see [21] or [22]), that $H_1 \in \mathcal{L}(\hat{\mathcal{X}}, \mathcal{X})$, $H_2 \in \mathcal{L}(\hat{\mathcal{X}}, \mathbb{R}^l)$, $R_1 \in \mathcal{L}(\mathcal{X})$ is (self-adjoint) nonnegative definite, and that R_2 is an $m \times m$ (symmetric) positive-definite matrix. We define $V_1 = H_1 H_1^*$ and $V_2 = H_2 H_2^*$, where $(\cdot)^*$ denotes adjoint, and assume for convenience that $H_1 H_2^* = 0$ and that V_2 is positive definite. The compensator is assumed to be of fixed, finite order n_c (i.e., $x_c(t) \in \mathbb{R}^{n_c}$) and that A_c, B_c , and C_c are matrices of appropriate dimension. For further details and discussion on the problem statement and the above assumptions, see [18].

We summarize here the primary result from [18] characterizing optimal fixed-finite-order controllers. For convenience define $\Sigma \triangleq B R_2^{-1} B^*$ and $\bar{\Sigma} \triangleq C^* V_2^{-1} C$. Also let I_{n_c} and $I_{\mathcal{X}}$ denote respectively the $n_c \times n_c$ identity matrix and the identity operator on \mathcal{X} .

Theorem 2.1. Let n_c be given and suppose there exists a controllable and observable n_c -th-order dynamic compensator (A_c, B_c, C_c) which minimizes J given by (2.5) and for which the closed-loop semigroup generated by

$$A \triangleq \begin{bmatrix} A & B C_c \\ B_c C & A_c \end{bmatrix}$$

is uniformly exponentially stable. Then there exist nonnegative-definite operators Q, P, \hat{Q}, \hat{P} on \mathcal{X} such that A_c, B_c , and C_c are given by

$$A_c = \Gamma(A - Q\bar{\Sigma} - \Sigma P)G^*, \quad (2.6)$$

$$B_c = \Gamma Q C^* V_2^{-1}, \quad (2.7)$$

$$C_c = -R_2^{-1} B^* P G^*, \quad (2.8)$$

where $Q : \text{Dom}(A^*) \rightarrow \text{Dom}(A)$, $P : \text{Dom}(A) \rightarrow \text{Dom}(A^*)$, $\hat{Q} : \mathcal{X} \rightarrow \text{Dom}(A)$, $\hat{P} : \mathcal{X} \rightarrow \text{Dom}(A^*)$, and $G, \Gamma \in \mathcal{L}(\mathcal{X}, \mathbb{R}^{n_c})$, and such that the following conditions are satisfied:

$$\text{rank } \hat{Q} = \text{rank } \hat{P} = \text{rank } \hat{Q}\hat{P} = n_c, \quad (2.9)$$

$$\hat{Q}\hat{P} = G^* M \Gamma, \quad \Gamma G^* = I_{n_c}, \quad (2.10)$$

for some $M \in \mathbb{R}^{n_c \times n_c}$,

$$0 = AQ + QA^* + V_1 - Q\bar{\Sigma}Q + \tau_{\perp}Q\bar{\Sigma}Q\tau_{\perp}^*, \quad (2.11)$$

$$0 = A^*P + PA + R_1 - P\Sigma P + \tau_{\perp}^*P\Sigma P\tau_{\perp}, \quad (2.12)$$

$$0 = (A - \Sigma P)\hat{Q} + \hat{Q}(A - \Sigma P)^* + Q\bar{\Sigma}Q - \tau_{\perp}Q\bar{\Sigma}Q\tau_{\perp}^*, \quad (2.13)$$

$$0 = (A - Q\bar{\Sigma})^*\hat{P} + \hat{P}(A - Q\bar{\Sigma}) + P\Sigma P - \tau_{\perp}^*P\Sigma P\tau_{\perp}, \quad (2.14)$$

where

$$\tau \triangleq G^* \Gamma, \quad \tau_{\perp} \triangleq I_X - \tau.$$

It is shown in [18] that the factorization (2.10) for the nonnegative-definite operators \hat{Q} and \hat{P} satisfying $\text{rank } \hat{Q}\hat{P} = n_c$ always exists and is unique except for a change of basis in \mathbb{R}^{n_c} . It is also shown in [18] that $\Gamma^* : \mathbb{R}^{n_c} \rightarrow \text{Dom}(A^*)$ so that the expression (2.6) is well defined.

Equations (2.11)–(2.14) are, in general, infinite-dimensional operator equations. To actually use them to compute the optimal fixed-order compensator, a finite-dimensional plant approximation is required. For each $N = 1, 2, \dots$, let \mathcal{X}^N denote a finite-dimensional subspace of \mathcal{X} and let $\mathcal{P}^N : \mathcal{X} \rightarrow \mathcal{X}^N$ denote the corresponding orthogonal projection of \mathcal{X} onto \mathcal{X}^N . Let $A^N \in \mathcal{L}(\mathcal{X}^N)$, $B^N \in \mathcal{L}(\mathbb{R}^m, \mathcal{X}^N)$, $C^N \in \mathcal{L}(\mathcal{X}^N, \mathbb{R}^l)$, $R_1^N \in \mathcal{L}(\mathcal{X}^N)$, and $V_1^N \in \mathcal{L}(\mathcal{X}^N)$. We consider the system (2.6)–(2.14) with the plant Π replaced by the plant Π^N given by

$$\Pi^N \triangleq \{A^N, B^N, C^N, R_1^N, R_2, V_1^N, V_2\}.$$

Typically, the operators B^N, C^N, R_1^N and V_1^N are chosen as $B^N = \mathcal{P}^N B$, $C^N = C\mathcal{P}^N$, $R_1^N = \mathcal{P}^N R_1$ and $V_1^N = \mathcal{P}^N V_1$ with the requirement that \mathcal{P}^N converge strongly to the identity I_X as $N \rightarrow \infty$. The operator A^N is chosen so that it and its adjoint satisfy the hypotheses of the Trotter-Kato semigroup approximation theorem (i.e., stability and consistency, see, for example, [23]). That is, A^N is chosen so that $\lim_{N \rightarrow \infty} T^N(t)\mathcal{P}^N\phi = T(t)\phi$, and $\lim_{N \rightarrow \infty} T^N(t)^*\mathcal{P}^N\phi = T(t)^*\phi$, uniformly in t for t in bounded intervals, for each $\phi \in \mathcal{X}$, where $T^N(t) = \exp(tA^N)$, $t \geq 0$. We shall say more about these choices for A^N, B^N, C^N, R_1^N , and V_1^N when we address convergence questions below.

Although with the plant Π^N equations (2.11)–(2.14) are finite dimensional, they are still operator equations. It is their *matrix* equivalents which are used in computations. Unless orthonormal bases are chosen for the subspaces \mathcal{X}^N (which is typically not the case in practice) some care must be taken to obtain the appropriate matrix system.

For each $N = 1, 2, \dots$, let $\{\phi_j^N\}_{j=1}^{k^N}$ be a basis for \mathcal{X}^N and choose the standard bases for all Euclidean spaces. For a linear operator L with domain and range \mathcal{X}^N or any Euclidean space, let $[L]$ denote its matrix representation with respect to the bases chosen above. Also, let Φ^N denote the k^N -square Gram matrix corresponding to the basis $\{\phi_j^N\}_{j=1}^{k^N}$; that is, $\Phi_{ij}^N = \langle \phi_i^N, \phi_j^N \rangle$, $i, j = 1, 2, \dots, k^N$. Noting that

$$\begin{aligned} [(A^N)^*] &= (\Phi^N)^{-1} [A^N]^T \Phi^N, & [(B^N)^*] &= [B^N]^T \Phi^N, & [(C^N)] &= (\Phi^N)^{-1} [C^N]^T, \\ [(\tau_\perp^N)^*] &= (\Phi^N)^{-1} [\tau_\perp^N]^T \Phi^N, & [\Sigma^N] &= [B^N] R_2^{-1} [B^N]^T \Phi^N, & [\bar{\Sigma}^N] &= (\Phi^N)^{-1} [C^N]^T V_2^{-1} [C^N], \end{aligned}$$

the matrix equivalents of the operator equations (2.11)–(2.14) become

$$\begin{aligned} 0 &= [A^N][Q^N] + [Q^N](\Phi^N)^{-1} [A^N]^T \Phi^N + [V_1^N] - [Q^N][\bar{\Sigma}^N][Q^N] \\ &\quad + [\tau_\perp^N][Q^N][\bar{\Sigma}^N][Q^N] - (\Phi^N)^{-1} [\tau_\perp^N]^T \Phi^N, \end{aligned} \quad (2.15)$$

$$\begin{aligned} 0 &= (\Phi^N)^{-1} [A^N]^T \Phi^N [P^N] + [P^N] - [A^N] + [R_1^N] - [P^N][\Sigma^N][P^N] \\ &\quad + (\Phi^N)^{-1} [\tau_\perp^N]^T \Phi^N [P^N] - [\Sigma^N][P^N][\tau_\perp^N], \end{aligned} \quad (2.16)$$

$$\begin{aligned} 0 &= ([A^N] - [\Sigma^N][P^N])[\hat{Q}^N] + [\hat{Q}^N](\Phi^N)^{-1} ([A^N] - [\Sigma^N][P^N])^T \Phi^N \\ &\quad + [Q^N][\bar{\Sigma}^N][Q^N] - [\tau_\perp^N][Q^N][\bar{\Sigma}^N][Q^N](\Phi^N)^{-1} - [\tau_\perp^N]^T \Phi^N, \end{aligned} \quad (2.17)$$

$$\begin{aligned} 0 &= (\Phi^N)^{-1} ([A^N] - [Q^N][\bar{\Sigma}^N])^T \Phi^N [\hat{P}^N] + [\hat{P}^N]([A^N] - [Q^N][\bar{\Sigma}^N]) \\ &\quad + [P^N][\Sigma^N][P^N] - (\Phi^N)^{-1} [\tau_\perp^N]^T \Phi^N [P^N][\Sigma^N][P^N][\tau_\perp^N]. \end{aligned} \quad (2.18)$$

Therefore, if we define the $k^N \times k^N$ nonnegative-definite matrices

$$\begin{aligned} Q_0^N &\triangleq [Q^N](\Phi^N)^{-1}, & P_0^N &\triangleq \Phi^N [P^N], \\ \hat{Q}_0^N &\triangleq [\hat{Q}^N](\Phi^N)^{-1}, & \hat{P}_0^N &\triangleq \Phi^N [\hat{P}^N], \\ V_0^N &\triangleq [V_1^N](\Phi^N)^{-1}, & R_0^N &\triangleq \Phi^N [R_1^N], \\ \Sigma_0^N &\triangleq [B^N] R_2^{-1} [B^N]^T, & \bar{\Sigma}_0^N &\triangleq [C^N]^T V_2^{-1} [C^N], \end{aligned}$$

we can solve the matrix optimal projection equations given in [14] corresponding to the matrix plant model

$$\Pi_0^N \triangleq \{[A^N], [B^N], [C^N], R_0^N, R_2, V_0^N, V_2\},$$

to obtain the matrices $Q_0^N, P_0^N, \hat{Q}_0^N$ and \hat{P}_0^N . The approximating optimal n_c th-order dynamic compensator $\{A_c^N, B_c^N, C_c^N\}$ is then given by

$$\begin{aligned} A_c^N &= \Gamma_0^N ([A^N] - Q_0^N \bar{\Sigma}_0^N - \Sigma_0^N P_0^N) (G_0^N)^T, \\ B_c^N &= \Gamma_0^N Q_0^N [C^N]^T V_2^{-1}, \\ C_c^N &= -R_2^{-1} [B^N]^T P_0^N (G_0^N)^T, \end{aligned}$$

where $\Gamma_0^N, G_0^N \in \mathbb{R}^{n_0 \times k^N}$, and $M^N \in \mathbb{R}^{n_0 \times n_0}$ satisfy

$$\hat{Q}_0^N \hat{P}_0^N = G_0^N M^N \Gamma_0^N, \quad \Gamma_0^N (G_0^N)^T = I_{n_0},$$

$$[r^N] = (G_0^N)^T \Gamma_0^N, \quad [\tau_\perp^N] = I_{k^N} - [r^N].$$

When an infinite-dimensional controller will suffice, $C_c = -R_2^{-1} B^* P \in \mathcal{L}(\mathcal{X}, \mathbb{R}^m)$ and $B_c = Q C^* V_2^{-1} \in \mathcal{L}(\mathbb{R}^\ell, \mathcal{X})$ are the usual infinite-dimensional LQG controller and observer gains (see [9]). The operators $P, Q \in \mathcal{L}(\mathcal{X})$ are the nonnegative-definite solutions to the two decoupled operator algebraic Riccati equations (2.11) and (2.12) with τ and τ_\perp formally set to $I_{\mathcal{X}}$ and 0, respectively. Since C_c has range in \mathbb{R}^m and B_c has domain \mathbb{R}^ℓ , there exist vectors $c_c = (c_c^1, \dots, c_c^m)^T \in \times_{j=1}^m \mathcal{X}$ and $b_c = (b_c^1, \dots, b_c^\ell)^T \in \times_{j=1}^\ell \mathcal{X}$ such that

$$[C_c x]_i = \langle c_c^i, x \rangle, \quad i = 1, 2, \dots, m, \quad x \in \mathcal{X},$$

and

$$B_c y = b_c^T y = \sum_{i=1}^\ell y_i b_c^i, \quad y = (y_1, \dots, y_\ell) \in \mathbb{R}^\ell.$$

The vectors c_c and b_c are referred to as the optimal LQG controller and observer functional gains respectively.

With regard to approximation for the full-order LQG problem, for each $N = 1, 2, \dots$ we take $n_c = k^N$. Then it is not difficult to show that

$$C_c^N [P^N x] = \langle c_c^N, x \rangle, \quad x \in \mathcal{X},$$

and

$$B_c^N y = (b_c^N)^T y, \quad y \in \mathbb{R}^\ell,$$

where $c_c^N \in \times_{j=1}^m \mathcal{X}^N$ and $b_c^N \in \times_{j=1}^\ell \mathcal{X}^N$ are given by $c_c^N = C_c^N (\Phi^N)^{-1} \phi^N$ and $b_c^N = (B_c^N)^T \phi^N$ respectively with $\phi^N = (\phi_1^N, \dots, \phi_{k^N}^N) \in \times_{j=1}^{k^N} \mathcal{X}^N$. The vectors c_c^N and b_c^N are referred to as the approximating optimal LQG controller and observer functional gains. To compute them we need only solve two standard decoupled matrix algebraic Riccati equations for the $k^N \times k^N$ nonnegative-definite matrices Q_0^N and P_0^N .

A rather complete convergence theory for LQG approximation can be found in [9]. Essentially, it is shown there that if the approximating subspaces \mathcal{X}^N are chosen so that the projections P^N

converge strongly to the identity as $N \rightarrow \infty$, the operators A^N, B^N, C^N, R_1^N , and V_1^N are chosen as was described above, and the operators Q^N and P^N are uniformly bounded in N , then Q^N and P^N converge weakly to Q and P , respectively as $N \rightarrow \infty$. This in turn implies that $C_c^N \rightarrow C_c$, strongly, $B_c^N \rightarrow B_c$, weakly, $c_c^N \rightarrow c_c$ and $b_c^N \rightarrow b_c$, weakly, and the closed-loop semigroup for the approximating optimal LQG compensator converges weakly to the closed-loop semigroup for the optimal infinite-dimensional LQG compensator, as $N \rightarrow \infty$. If, in addition, the operators $S^N(t) = T^N(t) + B^N C_c^N$ and $\hat{S}^N(t) = T^N(t) - B_c^N C^N$ are uniformly exponentially stable, uniformly in N , then $Q^N \rightarrow Q$ and $P^N \rightarrow P$, strongly, $C_c^N \rightarrow C_c$ and $B_c^N \rightarrow B_c$, in norm, $c_c^N \rightarrow c_c$ and $b_c^N \rightarrow b_c$, strongly, and the closed-loop semigroups converge strongly, as $N \rightarrow \infty$. If R_1^N and V_1^N are coercive, uniformly in N , then $S^N(t)$ and $\hat{S}^N(t)$ will be uniformly exponentially stable. If it is also true that R_1 and V_1 are trace class and $R_1^N P^N \rightarrow R_1$ and $V_1^N P^N \rightarrow V_1$ in trace norm then Q and P are trace class and $Q^N P^N \rightarrow Q$ and $P^N P^N \rightarrow P$ in trace norm as $N \rightarrow \infty$.

Returning to the fixed-finite-order case, we note that in general the approximating optimal projection equations may not possess a unique solution. However, in [19] it is shown for the finite-dimensional case that it is possible to obtain an upper bound for the number of stabilizing solutions. Using topological degree theory, the following result was obtained in [19].

Theorem 2.2. Consider the equations (2.11)–(2.14) with the infinite-dimensional plant Π replaced by the finite-dimensional plant Π^N . Let n_u denote the dimension of the unstable subspace of A^N and assume that $n_c \geq n_u$. Then in the class of nonnegative-definite operators $Q^N, P^N, \hat{Q}^N, \hat{P}^N$ on X^N satisfying $\text{rank } \hat{Q}^N = \text{rank } \hat{P}^N = \text{rank } \hat{Q}^N \hat{P}^N = n_c$, there exist at most

$$\begin{cases} \binom{\min(k^N, m, \ell) - n_u}{n_c - n_u}, & n_c \leq \min(k^N, m, \ell), \\ 1, & \text{otherwise,} \end{cases}$$

solutions of (2.11)–(2.14), each of which is stabilizing. If, in addition, the plant (A^N, B^N, C^N) is stabilizable by an n_c th-order controller, then there exists at least one stabilizing solution of (2.9)–(2.14).

Theorem 2.2 shows that while there may exist multiple solutions to the finite-dimensional optimal projection equations, in practice this number can be quite small. For example, if $n_c \geq n_u$ and the system is either single input ($m = 1$) or single output ($\ell = 1$) then there exists at most one solution to (2.9)–(2.14) for the plant Π^N . The existence of at least one stabilizing solution of course depends upon whether or not the plant is stabilizable by an n_c th-order controller

(for relevant results, see [24]). Finally, while it may be possible to stabilize a plant with $n_c < n_u$, this case lies outside the scope of the analysis given in [19].

3. Examples and Numerical Results

We first consider the one-dimensional, single-input/ single-output, parabolic (heat/diffusion) control system with Dirichlet boundary conditions given by

$$\frac{\partial v}{\partial t}(t, \eta) = a \frac{\partial^2 v}{\partial \eta^2}(t, \eta) + b(\eta)u(t) + h_1(\eta)w_1(t, \eta), \quad 0 < \eta < 1, \quad t > 0, \quad (3.1)$$

$$v(t, 0) = 0, \quad v(t, 1) = 0, \quad t > 0, \quad (3.2)$$

$$y(t) = \int_0^1 c(\eta)v(t, \eta)d\eta + h_2w_2(t), \quad t > 0, \quad (3.3)$$

where $a > 0$, and $b(\cdot)$ and $c(\cdot)$ are given by

$$b(\eta) = \begin{cases} \frac{1}{\beta_2 - \beta_1}, & \beta_1 \leq \eta \leq \beta_2, \\ 0, & \text{elsewhere,} \end{cases}$$

and

$$c(\eta) = \begin{cases} \frac{1}{\gamma_2 - \gamma_1}, & \gamma_1 \leq \eta \leq \gamma_2, \\ 0, & \text{elsewhere,} \end{cases}$$

with $0 \leq \beta_1 < \beta_2 \leq 1$ and $0 \leq \gamma_1 < \gamma_2 \leq 1$. In (3.1) and (3.3), $h(\cdot) \in L_\infty(0, 1)$, $w_1(t, \cdot) \in L_2(0, 1)$, a.a. $t \in [0, \infty)$, (see [22], p. 314), h_2 is a nonzero constant and $w_2(\cdot)$ is unit-intensity white noise.

To rewrite (3.1)–(3.3) in the form (2.1), (2.2), in the usual way we take $\mathcal{X} = L_2(0, 1)$ endowed with the standard L_2 inner product, let $x(t) = v(t, \cdot)$, $t \geq 0$, define $A : \text{Dom}(A) \subset \mathcal{X} \rightarrow \mathcal{X}$ by $A\phi = aD^2\phi$ for $\phi \in \text{Dom } A \triangleq H^2(0, 1) \cap H_0^1(0, 1)$, and define $B \in \mathcal{L}(\mathbb{R}^1, \mathcal{X})$ and $C \in \mathcal{L}(\mathcal{X}, \mathbb{R}^1)$ by $Bu = b(\cdot)u$ for $u \in \mathbb{R}^1$, and $C\phi = \int_0^1 c(\eta)\phi(\eta)d\eta$, for $\phi \in L_2(0, 1)$, respectively. Furthermore, let $\hat{\mathcal{X}} \triangleq L_2(0, 1) \times \mathbb{R}$, set $w(t) \triangleq (w_1(t, \cdot), w_2(t)) \in \hat{\mathcal{X}}$, and define $H_1 \in \mathcal{L}(\hat{\mathcal{X}}, \mathcal{X})$ and $H_2 \in \mathcal{L}(\hat{\mathcal{X}}, \mathbb{R}^1)$ by $H_1z = h_1(\cdot)z_1$ and $H_2z = h_2z_2$ for $z = (z_1, z_2) \in \hat{\mathcal{X}}$.

It is well known (see, for example, [23]) that A is closed, densely defined, and negative definite. Furthermore, A is the infinitesimal generator of a uniformly exponentially stable, analytic (abstract parabolic) semigroup $\{T(t) : t \geq 0\}$ of bounded, self-adjoint linear operators on \mathcal{X} .

We consider linear spline-based Ritz-Galerkin approximation for the open-loop system. For each $N = 2, 3, \dots$, let $\{\phi_j^N\}_{j=1}^{N-1}$ be the linear spline ("hat") functions defined on the interval $[0, 1]$

with respect to the uniform partition $\{0, \frac{1}{N}, \frac{2}{N}, \dots, 1\}$, i.e.,

$$\phi_j^N(\eta) = \begin{cases} N\eta - j + 1, & \eta \in [\frac{j-1}{N}, \frac{j}{N}), \\ j + 1 - N\eta, & \eta \in [\frac{j}{N}, \frac{j+1}{N}), \\ 0, & \text{elsewhere on } [0,1], \end{cases}$$

$j = 1, 2, \dots, N-1$. Set $\mathcal{X}^N = \text{span}\{\phi_j^N\}_{j=1}^{N-1}$ and note that $k^N = \dim \mathcal{X}^N = N-1$, and $\mathcal{X}^N \subset H_0^1(0,1)$ for all N . If $\mathcal{P}^N : \mathcal{X} \rightarrow \mathcal{X}^N$ denotes the orthogonal projection of \mathcal{X} onto \mathcal{X}^N , then standard convergence estimates for interpolatory splines (see [25]) can be used to show that $\lim_{N \rightarrow \infty} \mathcal{P}^N \phi = \phi$ in $L_2(0,1)$ for $\phi \in L_2(0,1)$.

There are two equivalent ways to obtain an operator representation for the usual Ritz-Galerkin approximation to A . First, A can be extended to a bounded linear operator from $H_0^1(0,1)$ onto its dual, $H^{-1}(0,1)$, via

$$(A\phi)(\psi) = -a\langle D\phi, D\psi \rangle, \quad \phi, \psi \in H_0^1(0,1). \quad (3.4)$$

Since $\mathcal{X}^N \subset H_0^1(0,1)$ for all $N = 2, 3, \dots$, we define $A^N \in \mathcal{L}(\mathcal{X}^N)$ by $A^N \phi^N = A\phi^N$, $\phi^N \in \mathcal{X}^N$, with $A\phi^N \in H^{-1}(0,1)$ considered to be a linear functional on \mathcal{X}^N . From the Riesz Representation theorem we obtain $A^N \phi^N = \psi^N$ where ψ^N is that element in \mathcal{X}^N which satisfies $(A^N \phi^N)(\chi^N) = -a\langle D\phi^N, D\chi^N \rangle = \langle \psi^N, \chi^N \rangle$.

Alternatively and equivalently, by using the fact that A is self-adjoint, we can define A^N as follows. Let $\mathcal{P}_1^N : H_0^1(0,1) \rightarrow \mathcal{X}^N$ denote the orthogonal projection of the Hilbert space $H_0^1(0,1)$ onto \mathcal{X}^N . Using the definition (3.4), it is not difficult to show that $-A \in \mathcal{L}(H_0^1(0,1), H^{-1}(0,1))$ is coercive and, therefore, that $A^{-1} : H^{-1}(0,1) \rightarrow H_0^1(0,1)$ exists and is bounded. We then define $A^N \in \mathcal{L}(\mathcal{X}^N)$ to be the inverse of the operator given by $(A^N)^{-1} = \mathcal{P}_1^N A^{-1} |_{\mathcal{X}^N}$.

Using either definition, it is easily argued that A^N is well defined, self-adjoint, and is the infinitesimal generator of a uniformly exponentially stable (uniformly in N) semigroup, $T^N(t) = \exp(tA^N)$, $t \geq 0$, of bounded linear operators on \mathcal{X}^N . Also, using the approximation properties of splines, it is not difficult to show that $\lim_{N \rightarrow \infty} (A^N)^{-1} \mathcal{P}^N \phi = A^{-1} \phi$, $\phi \in \mathcal{X}$. Consequently, the hypotheses of the Trotter-Kato theorem (see [23]) are satisfied and we have $\lim_{N \rightarrow \infty} T^N(t) \mathcal{P}^N \phi = T(t)\phi$ and $\lim_{N \rightarrow \infty} T^N(t)^* \mathcal{P}^N \phi = T(t)^* \phi$, $\phi \in \mathcal{X}$, uniformly in t for t in bounded intervals. A detailed discussion of the results just outlined can be found in [8].

We define $B^N = \mathcal{P}^N B$ and $C^N = C \mathcal{P}^N$, from which it immediately follows that $\lim_{N \rightarrow \infty} B^N = B$ and $\lim_{N \rightarrow \infty} C^N = C$ in norm and similarly for their adjoints. For the example we shall consider

here, we have chosen $R_1 = r_1 I_X$, $R_2 = r_2 I_m$, with $r_1, r_2 > 0$. Setting $h_1(\eta) = v_1^{\frac{1}{2}}$, $0 < \eta < 1$, and $h_2 = v_2^{\frac{1}{2}}$ with $v_1, v_2 > 0$, we obtain $V_1 = v_1 I_X$ and $V_2 = v_2$. We then take $R_1^N = P^N R_1$ and $V_1^N = P^N V_1$. For the LQG problem, the open-loop uniform exponential stability of both the infinite-dimensional system and the approximating systems is sufficient to conclude the strong convergence of the approximating Riccati operators to the solutions of the infinite-dimensional Riccati equations, the uniform norm convergence of the approximating controller and observer gains, and the strong convergence of the functional gains, as $N \rightarrow \infty$.

Since the basis elements $\{\phi_j^N\}_{j=1}^{N-1}$ are piecewise linear with respect to the uniform mesh $\{0, \frac{1}{N}, \frac{2}{N}, \dots, 1\}$ on $[0, 1]$, the equivalent matrix representations for the operators defined above can be computed directly and in closed form. The Gram matrix $\Phi_{ij}^N = \langle \phi_i^N, \phi_j^N \rangle$, $i, j = 1, 2, \dots, N-1$ is given by $\Phi^N = \frac{1}{N} \text{Tridiag}\{\frac{1}{6}, \frac{2}{3}, \frac{1}{6}\}$, and if we define the generalized stiffness matrix Ψ^N by $\Psi_{ij}^N = -a \langle D\phi_i^N, D\phi_j^N \rangle$, $i, j = 1, 2, \dots, N-1$, then $\Psi^N = aN \text{Tridiag}\{1, -2, 1\}$. It follows that $[A^N] = (\Phi^N)^{-1} \Psi^N$, $[B^N] = (\Phi^N)^{-1} b^N$, $[C^N] = c^N$, with $b_i^N = \langle b, \phi_i^N \rangle = \frac{1}{\beta_2 - \beta_1} \int_{\beta_1}^{\beta_2} \phi_i^N(\eta) d\eta$, and $c_i^N = \langle c, \phi_i^N \rangle = \frac{1}{\gamma_2 - \gamma_1} \int_{\gamma_1}^{\gamma_2} \phi_i^N(\eta) d\eta$, $i = 1, 2, \dots, N-1$, and that $R_0^N = r_1 \Phi^N$ and $V_0^N = v_1 (\Phi^N)^{-1}$.

For our numerical study we set $a = 1$, $\beta_1 = .75 - .03\sqrt{2}$, $\beta_2 = .75 + .04\sqrt{2}$, $\gamma_1 = .25 - .04\sqrt{2}$, $\gamma_2 = .25 + .03\sqrt{2}$, $r_1 = v_1 = 1$, $r_2 = v_2 = 10^{-4}$, $h_1(\eta) \equiv 1$, and used our technique to compute approximating optimal LQG (i.e., $n_c = N-1$) and 1st order (i.e., $n_c = 1$) compensators for various values of N . The open-loop stability of system (3.1)–(3.3) and the approximating systems imply that the finite-dimensional approximating optimal projection equations have a solution. Theorem 2.2 on the other hand, with $n_u = 0$ and $n_c = 1$ or $n_c = N-1$, implies that they have at most one solution. Consequently, the system of equations (2.11)–(2.14) with the plants Π^N admits a unique solution.

The optimal projection equations (2.11)–(2.14) were solved using the homotopic continuation algorithm described in [19]. It is shown in [19] that the operation count for the algorithm is proportional to $p(2n^3 + (m + \ell)n^2 + (m + \ell)^3 n_c^3)$ where p is the number of integration steps and n is the dimension of the finite-dimensional plant. This is competitive with the operation count for the Hamiltonian solution of the standard Riccati equations which is $O(16n^3)$ for LQG. Also, note that the computational burden for the solution of the optimal projection equations decreases with n_c .

Since $m = \ell = 1$ in the LQG case, the optimal functional observer and feedback control gains b_c and c_c and the approximating gains b_c^N and c_c^N , are all simply L_2 functions with b_c^N and c_c^N elements in \mathcal{X}^N . We plot the functions b_c^N and c_c^N we obtained for various values of N respectively

in Figures 3.1 and 3.2 below. That convergence is indeed achieved can immediately be observed in the figures.

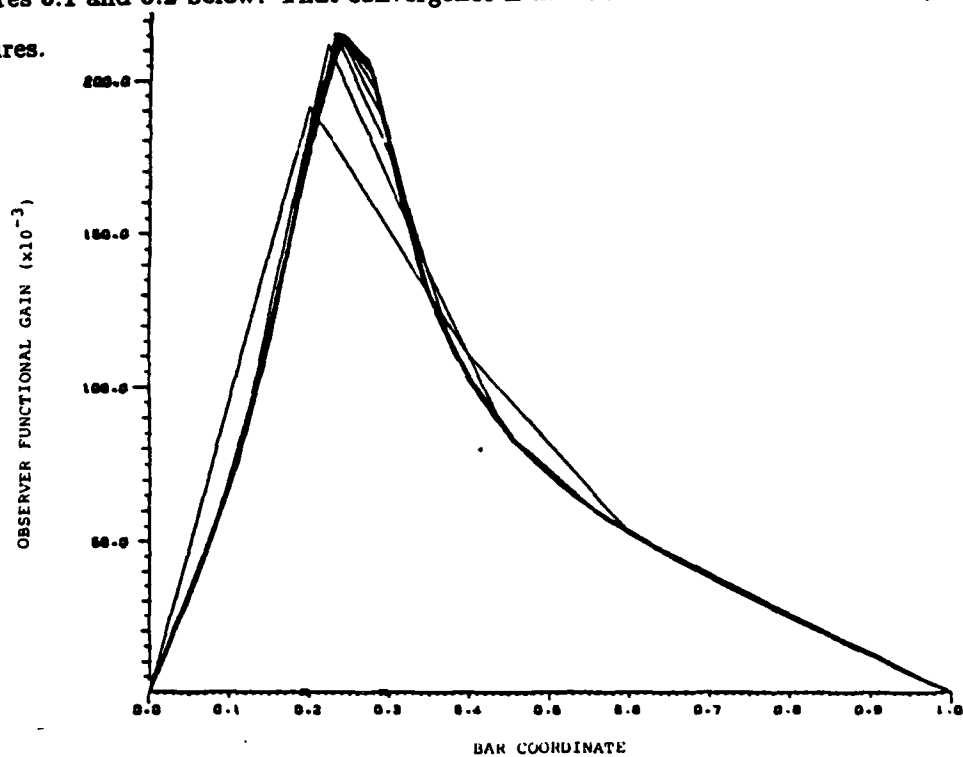


Figure 3.1

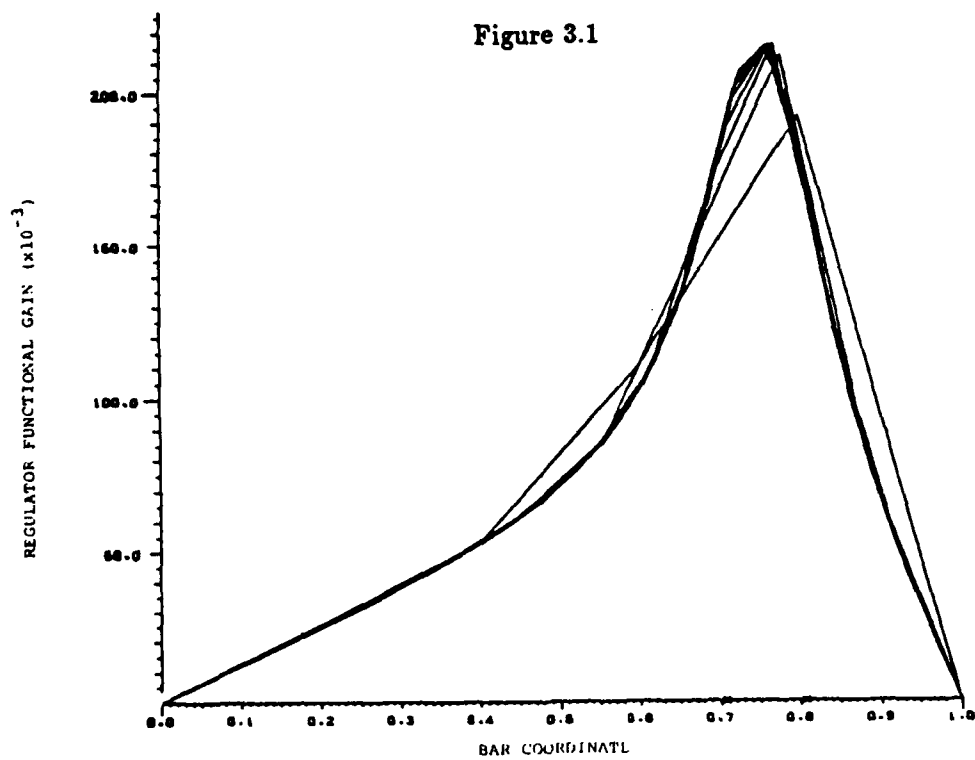


Figure 3.2

In the fixed-order case with $n_c = 1$, the compensator gains A_c , B_c , and C_c are all scalars. Also, for a first-order controller there are only two independent parameters, A_c and $B_c C_c$. In Table 3.1 below we give the values we obtained for A_c^N and $B_c^N C_c^N$ for various values of N . Once again, it is clear that the gains are converging as N increases. In addition, in Table 3.1 we provide the closed-loop costs J_{LQG}^N and J_{FO}^N for the LQG and first-order controllers. These closed-loop costs were evaluated using a 64th-order modal approximation to the infinite-dimensional system. For all values of N the performance of the fixed-order compensator was within 2% of the corresponding LQG controller. Thus, for example, the replacement of a 32nd-order approximating optimal LQG controller by an approximating optimal first-order controller will yield considerable implementation simplification with only minor performance degradation. Note that for the example we consider here, it is possible to compute the open-loop cost for the infinite-dimensional system in closed form. We have

$$\begin{aligned} J_{OL} &= \text{tr} \int_0^\infty V_1 T^*(t) R_1 T(t) dt = v_1 r_1 \text{tr} \int_0^\infty T(t)^2 dt \\ &= v_1 r_1 \sum_{n=1}^\infty \int_0^\infty e^{-2n^2 \pi^2 a t} dt = \frac{v_1 r_1}{2\pi^2 a} \sum_{n=1}^\infty \frac{1}{n^2} \\ &= \frac{v_1 r_1}{12a} = \frac{1}{12} \approx .08333. \end{aligned}$$

Finally, for comparison purposes, we tried applying balancing techniques to the LQG controllers to reduce their order. However, with $n_c = 1$, such controllers were found to be destabilizing. Based upon the results in [13], this was not unexpected.

| N | A_c^N | $B_c^N C_c^N$ | J_{LQG}^N | J_{FO}^N |
|-----|---------|---------------|-------------|------------|
| 4 | -687.6 | 5470 | .06999 | .07014 |
| 8 | -720.9 | 5231 | .06870 | .06993 |
| 12 | -730.9 | 5182 | .06872 | .06991 |
| 16 | -734.3 | 5145 | .06874 | .06990 |
| 20 | -738.0 | 5127 | .06875 | .06990 |
| 24 | -737.6 | 5108 | .06876 | .06990 |
| 28 | -739.8 | 5109 | .06876 | .06990 |
| 32 | -738.7 | 5099 | .06877 | .06990 |

Table 3.1

As a second example we consider the one-dimensional, single-input, single-output hereditary control system given by

$$\dot{v}(t) = a_0 v(t) + a_1 v(t - \rho) + b_0 u(t) + h_1 w(t), \quad t > 0, \quad (3.5)$$

$$y(t) = c_0 v(t) + h_2 w(t), \quad t > 0, \quad (3.6)$$

where $a_0, a_1, b_0, c_0, h_1, h_2, \rho \in \mathbb{R}^1$ with $h_2 \neq 0$, and w is a unit-intensity white noise process. To rewrite (3.5), (3.6) in the form (2.1), (2.2), we take $\mathcal{X} = \mathbb{R}^1 \times L_2(-\rho, 0)$ endowed with the usual product space inner product, $\langle (\eta, \phi), (\xi, \psi) \rangle = \eta\xi + \int_{-\rho}^0 \phi\psi$, and let $x(t) = (v(t), v_t)$, $t \geq 0$, where for $t \geq 0$, $v_t \in L_2(-\rho, 0)$ is given by $v_t(\theta) = v(t + \theta)$, $-\rho \leq \theta \leq 0$. Define $A: \text{Dom}(A) \subset \mathcal{X} \rightarrow \mathcal{X}$ by $A\hat{\phi} = (a_0\phi(0) + a_1\phi(-\rho), D\phi)$ for $\hat{\phi} = (\phi(0), \phi) \in \text{Dom}(A) \triangleq \{(\xi, \psi) \in \mathcal{X} : \psi \in H^1(-\rho, 0), \psi(0) = \xi\}$, and let $B \in \mathcal{L}(\mathbb{R}^1, \mathcal{X})$ and $C \in \mathcal{L}(\mathcal{X}, \mathbb{R}^1)$ be given by $Bu = (b_0 u, 0)$ and $C(\eta, \phi) = c_0 \eta$, respectively. Let $\hat{\mathcal{X}} = \mathbb{R}^1$ and define $H_1 \in \mathcal{L}(\hat{\mathcal{X}}, \mathcal{X})$ and $H_2 \in \mathcal{L}(\hat{\mathcal{X}}, \mathbb{R}^1)$ by $H_1 z = (h_1 z, 0)$ and $H_2 z = h_2 z$, for $z \in \mathbb{R}^1$.

The operator A is densely defined and is the infinitesimal generator of a C_0 semigroup $\{T(t): t \geq 0\}$ of bounded linear operators on \mathcal{X} with $T(t)(\eta, \phi) = (v(t; \eta, \phi), v_t(\eta, \phi))$, $t \geq 0$, where $v(\cdot; \eta, \phi)$ is the unique solution to (3.5) with $b_0 = h_1 = 0$, and initial conditions $v(0) = \eta$, $v_0 = \phi$. We take $R_1 \in \mathcal{L}(\mathcal{X})$ and $R_2 \in \mathcal{L}(\mathbb{R}^1)$ to be $R_1(\eta, \phi) = (r_1 \eta, 0)$ and $R_2 u = r_2 u$, respectively, with $r_1, r_2 > 0$. The definitions of H_1 and H_2 given above imply that $V_1 \in \mathcal{L}(\mathcal{X})$ and $V_2 \in \mathcal{L}(\mathbb{R}^1)$ are given by $V_1(\eta, \phi) = (h_1^2 \eta, 0)$ and $V_2 z = h_2^2 z$, for $(\eta, \phi) \in \mathcal{X}$ and $z \in \mathbb{R}^1$.

We employ an approximation scheme recently proposed by Ito and Kappel in [26]. We briefly outline it here; a more detailed discussion can be found in [26]. For each $N = 1, 2, \dots$ let $\chi_j^N \in L_2(-\rho, 0)$ denote the characteristic function for the interval $[-j\rho/N, -(j-1)\rho/N]$, $j = 1, 2, \dots, N$, and let \mathcal{X}^N be the $(N+1)$ -dimensional subspace of \mathcal{X} defined by

$$\mathcal{X}^N = \text{span}\{(1, 0), (0, \chi_1^N), \dots, (0, \chi_N^N)\}.$$

Let $\mathcal{P}^N: \mathcal{X} \rightarrow \mathcal{X}^N$ denote the orthogonal projection of \mathcal{X} onto \mathcal{X}^N . Let $\{\phi_j^N\}_{j=0}^N$ denote the linear B-spline functions defined on the interval $[-\rho, 0]$ with respect to the uniform mesh $[-\rho, \dots, -\rho/N, 0]$, and set $\mathcal{X}_1^N = \text{span}\{(\phi_j^N(0), \phi_j^N)\}_{j=0}^N$. Then \mathcal{X}_1^N is an $(N+1)$ -dimensional subspace of $\text{Dom}(A)$ and it is not difficult to demonstrate that the restriction of \mathcal{P}^N to \mathcal{X}_1^N is a bijection onto \mathcal{X}^N . Using the fact that A restricted to \mathcal{X}_1^N has range in \mathcal{X}^N , we define $A^N \in \mathcal{L}(\mathcal{X}^N)$ by $A^N = A(\mathcal{P}^N)^{-1}$, and set $T^N(t) = \exp(A^N t)$, $t \geq 0$. Noting that $R(B) \subset \mathcal{X}^N$, we take $B^N \in \mathcal{L}(\mathbb{R}^1, \mathcal{X}^N)$ to be given by $B^N = B$. Similarly, we take $R_1^N = R_1$ and $V_1^N = V_1$. We set $C^N = C$.

It is shown in [26] that $P^N(\eta, \phi) \rightarrow (\eta, \phi)$, $T^N(t)P^N(\eta, \phi) \rightarrow T(t)(\eta, \phi)$, and $T^N(t)^*P^N(\eta, \phi) \rightarrow T(t)^*(\eta, \phi)$ for $(\eta, \phi) \in \mathcal{X}$ as $N \rightarrow \infty$, uniformly in t , for t in bounded subsets of $[0, \infty)$. It then follows that $\lim_{N \rightarrow \infty} B^N = B$ and $\lim_{N \rightarrow \infty} C^N P^N = C$, in norm.

For the LQG (full-order) problem, the optimal functional observer and feedback control gains b_c and c_c are of the form $b_c = (\beta_0, \beta_1)$ and $c_c = (\gamma_0, \gamma_1)$ with $\beta_0, \gamma_0 \in \mathbb{R}^1$, and $\beta_1, \gamma_1 \in L_2(-\rho, 0)$. The approximating gains are of the form $b_c^N = (\beta_0^N, \beta_1^N)$ and $c_c^N = (\gamma_0^N, \gamma_1^N)$ with $\beta_0^N, \gamma_0^N \in \mathbb{R}^1$ and $\beta_1^N, \gamma_1^N \in \text{span} \{\chi_j^N\}_{j=1}^N$. Since we are treating a one-dimensional example, if $b_0 \neq 0$, the theory in [26] implies that $\beta_0^N \rightarrow \beta_0$ and $\gamma_0^N \rightarrow \gamma_0$ in \mathbb{R}^1 , and $\beta_1^N \rightarrow \beta_1$, and $\gamma_1^N \rightarrow \gamma_1$ in $L_2(-\rho, 0)$, as $N \rightarrow \infty$.

Once again, as in the first example, matrix representations for the operators A^N , B^N , C^N , R_1^N , and V_1^N are not difficult to compute in closed form. Indeed, the $(N+1) \times (N+1)$ matrix representation for the bijection $P^N: \mathcal{X}_1^N \rightarrow \mathcal{X}_1$ is given by

$$[P^N] = \begin{bmatrix} 1 & 0 & & & 0 \\ \frac{1}{2} & \frac{1}{2} & 0 & & \\ 0 & \frac{1}{2} & \frac{1}{2} & 0 & \\ & & \ddots & \ddots & \\ 0 & & & 0 & \frac{1}{2} & \frac{1}{2} & 0 \\ & & & & 0 & \frac{1}{2} & \frac{1}{2} \end{bmatrix}.$$

Then $[A^N] = [K^N][P^N]^{-1}$, where

$$[K^N] = \begin{bmatrix} a_0 & 0 & & & 0 & & a_1 \\ \frac{N}{\rho} & -\frac{N}{\rho} & 0 & & & & \\ 0 & \frac{N}{\rho} & -\frac{N}{\rho} & 0 & & & \\ & & \ddots & \ddots & \ddots & & \\ & & & 0 & \frac{N}{\rho} & -\frac{N}{\rho} & 0 \\ & & & & 0 & \frac{N}{\rho} & -\frac{N}{\rho} \end{bmatrix}.$$

We have the $(N+1) \times 1$ matrix $[B^N] = [b_0 \ 0 \ \dots \ 0]^T$ and the $1 \times (N+1)$ matrix $[C^N] = [c_0 \ 0 \ \dots \ 0]$, while $[R_1^N] = r_1[\mathcal{M}^N]$ and $[V_1^N] = h_1^2[\mathcal{M}^N]$ where the $(N+1) \times (N+1)$ matrix $[\mathcal{M}^N]$ is given by

$$[\mathcal{M}^N] = \begin{bmatrix} 1 & 0 & & & 0 \\ 0 & & & & 0 \\ & & \ddots & & \\ 0 & & & & 0 \end{bmatrix}.$$

We set $a_0 = a_1 = b_0 = c_0 = r_1 = h_1 = \rho = 1$, $r_2 = .1$, and $h_2 = \sqrt{.1}$ and computed approximating optimal LQG (i.e., $n_c = N+1$) and first-order (i.e., $n_c = 1$) compensators for

$N = 8, 16, 24$, and 32 . The optimal LQG observer gains are given in Table 3.3 and Figure 3.3; the control gains are given in Table 3.4 and Figure 3.4. The first 23 open-loop poles of the system (see [27]) are given in Table 3.2. The approximating first-order compensator gains along with the corresponding and LQG closed-loop costs are given in Table 3.5 below. These costs were computed using an evaluation model obtained by setting $N = 64$. Note that the performance of the first-order controllers is within 10% of the performance of the LQG controllers. Once again it is clear that convergence is achieved.

| |
|---------------------------|
| 1.278465 |
| -1.588317 \pm 4.155305i |
| -2.417631 \pm 10.68603i |
| -2.861502 \pm 17.05611i |
| -3.167754 \pm 23.38558i |
| -3.401945 \pm 29.69798i |
| -3.591627 \pm 36.00146i |
| -3.751047 \pm 42.29965i |
| -3.888543 \pm 48.59442i |
| -4.009422 \pm 54.88686i |
| -4.117267 \pm 61.17761i |
| -4.214618 \pm 67.46710i |

Table 3.2

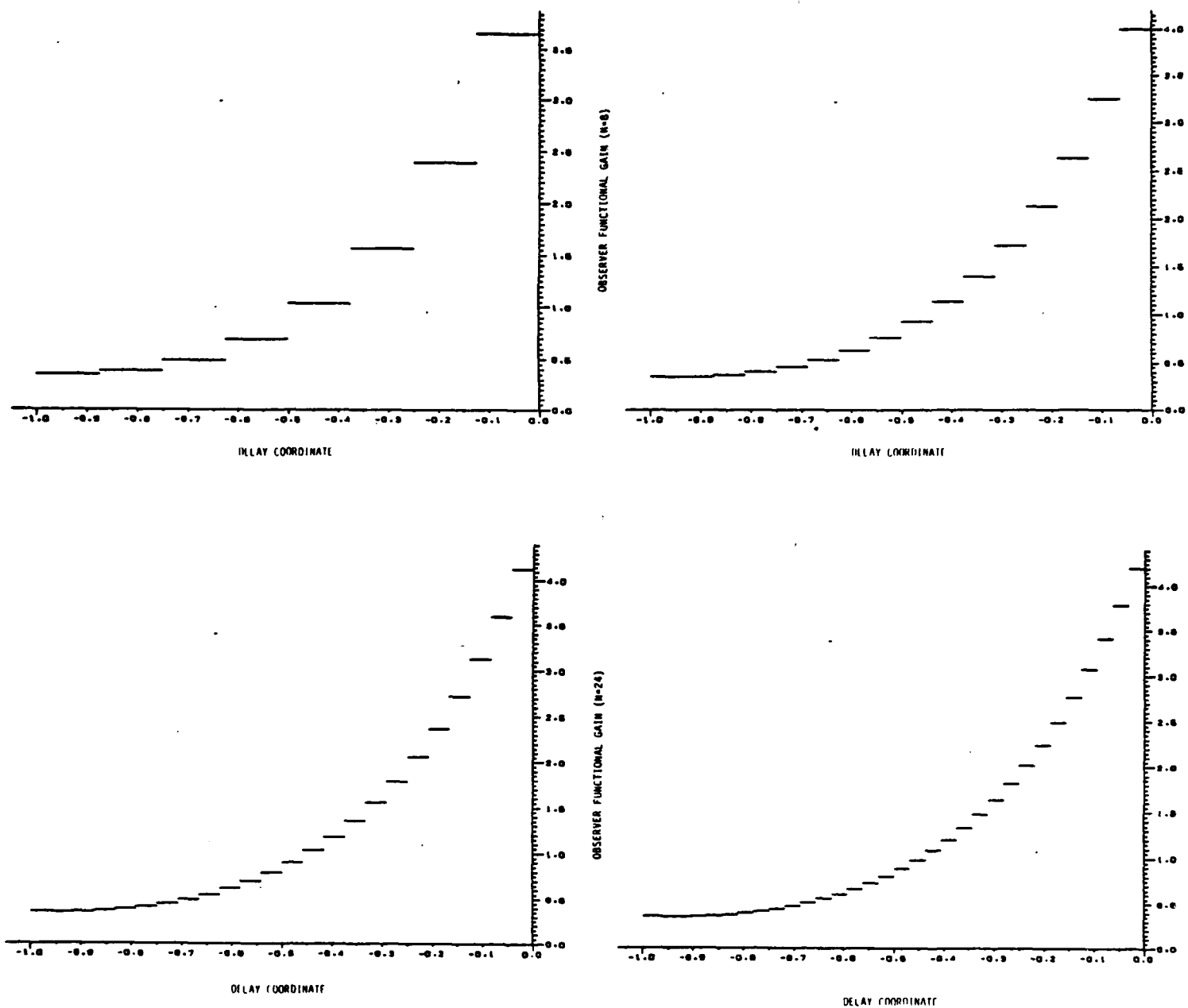


Figure 3.3

| N | 8 | 16 | 24 | 32 |
|-------------|--------|--------|--------|--------|
| β_0^N | 4.4213 | 4.4229 | 4.4233 | 4.4234 |

Table 3.3

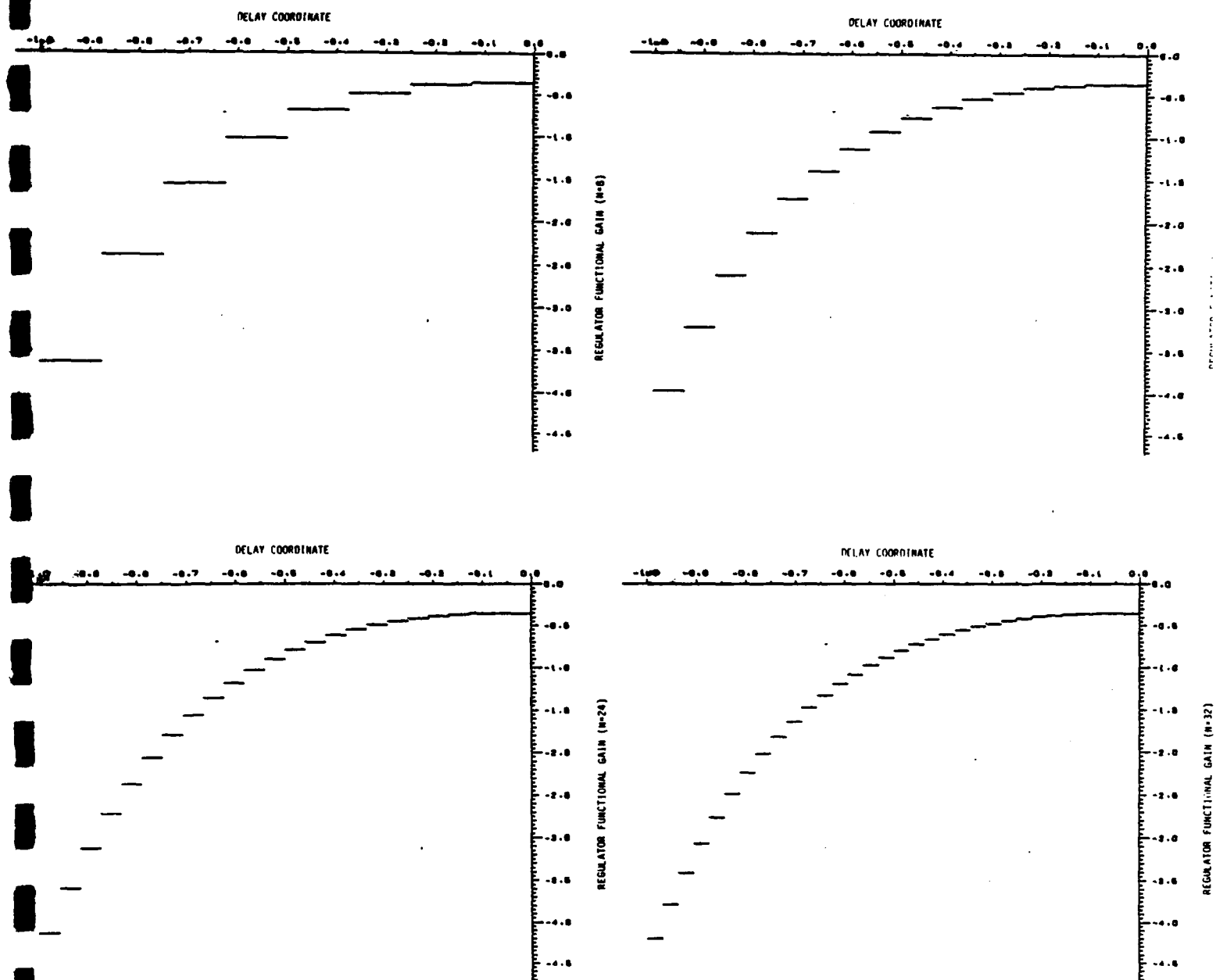


Figure 3.4

| N | 8 | 16 | 24 | 32 |
|--------------|---------|---------|---------|---------|
| γ_0^N | -4.4213 | -4.4229 | -4.4233 | -4.4234 |

Table 3.4

| N | A_c^N | $B_c^N C_c^N$ | J_{LQG}^N | J_{FO}^N |
|-----|---------|---------------|-------------|------------|
| 8 | -4.835 | -16.057 | 1.4042 | 1.5221 |
| 16 | -4.936 | -16.343 | 1.403877 | 1.5298 |
| 24 | -4.959 | -16.378 | 1.403856 | 1.5309 |
| 32 | -4.962 | -16.404 | 1.403852 | 1.5317 |

Table 3.5

4. Summary and Concluding Remarks

We have proposed an approximation technique for computing optimal fixed-order compensators for distributed parameter systems. Our approach involves using the optimal projection theory for infinite-dimensional systems (which characterizes the optimal fixed-order compensator) developed in [18] in conjunction with finite-dimensional approximation of the infinite-dimensional plant. We demonstrated the feasibility of our approach with two examples wherein we used spline-based Ritz-Galerkin finite element schemes to compute approximating optimal first-order controllers for one-dimensional, single-input/output, parabolic (heat/diffusion) and hereditary control systems. Our numerical studies indicate that convergence of the compensator gains is achieved and that using the first-order controller would lead to only minimal performance degradation over a standard LQG compensator while yielding significant implementation simplification.

At this point one is led naturally to ask the question of whether or not a satisfactory convergence theory could be developed. We are working on this at present and expect that such a theory would conform closely in form and spirit to the convergence results for LQG approximation found in [9] and [10] and outlined in Section 2 above. We also intend to consider our approximation ideas in the context of discrete-time or sampled-data systems, and for continuous-time systems involving unbounded input and/or output (for example, boundary control systems), and systems with control or measurement delays, see [11],[12]). Finally, we intend to investigate the application of our approximation framework to other infinite-dimensional control systems, in particular the vibration control of flexible structures (i.e., second-order systems such as wave, beam, or plate equations).

Acknowledgment. The authors gratefully acknowledge Mr. S. W. Greeley, Mr. S. Richter, and Mr. A. Daubendiek of Harris Corp. for carrying out the numerical computations reported in Section 3. We also wish to thank Ms. J. M. Straehla, also of Harris Corp., for preparing the original manuscript in \TeX .

References

1. Banks, H. T. and J. A. Burns, "Hereditary control problems: Numerical methods based on averaging approximations," *SIAM J. Control and Opt.*, 16, 169-208 (1978).
2. Gibson, J. S., "The Riccati integral equations for optimal control problems on Hilbert spaces," *SIAM J. Control and Opt.*, 17, 537-565 (1979).
3. Gibson, J. S., "An analysis of optimal modal regulation: Convergence and stability," *SIAM J. Control and Opt.*, 19, 686-707 (1981).
4. Kunisch, K., "Approximation schemes for the linear quadratic optimal control problem associated with delay equations," *SIAM J. Control and Opt.*, 20, 506-540 (1982).
5. Gibson, J. S., "Linear quadratic optimal control of hereditary differential systems: Infinite dimensional Riccati equations and numerical approximations," *SIAM J. Control and Opt.*, 21, 95-139 (1983).
6. Banks, H. T., I. G. Rosen, and K. Ito, "A spline based technique for computing Riccati operators and feedback controls in regulator problems for delay equations," *SIAM J. Sci. Stat. Comput.*, 5, 830-855 (1984).
7. Banks, H. T. and K. Kunisch, "The linear regulator problem for parabolic systems," *SIAM J. Control and Opt.*, 22, 684-698 (1984).
8. Gibson, J. S. and I. G. Rosen, "Numerical approximation for the infinite dimensional discrete time optimal linear quadratic regulator problem," *SIAM J. Control and Opt.*, 26, pp. 428-451 (1988).
9. Gibson, J. S. and A. D. Adamian, "Approximation theory for LQG optimal control of flexible structures," Technical Report, Dept. Mech., Aero., Nuc. Eng., University of California, Los Angeles, CA, 1986.
10. Gibson, J. S. and I. G. Rosen, "Computational methods for optimal linear quadratic compensators for infinite dimensional discrete time systems," *Proceedings of 3rd International Conference on Distributed Parameter Systems*, Vorau, Austria, July 6-12, 1986, Lecture Notes in Control and Information Sciences No. 102, Springer-Verlag, New York, 1987, 120-135.
11. Gibson, J. S. and I. G. Rosen, "Approximation of discrete time LQG compensators for distributed systems with boundary input and unbounded measurement," *Automatica*, Vol. 23, 1988.
12. Gibson, J. S. and I. G. Rosen, "Approximation in discrete time boundary control of flexible structures," *Proceedings 26th IEEE Conference on Decision and Control*, 535-540, Los Angeles, CA, Dec. 9-11, 1987.
13. Liu, Y. and B. D. O. Anderson, "Controller reduction via stable factorization and balancing," *Int. J. Control*, 44, 507-531 (1986).
14. Hyland, D. C. and D. S. Bernstein, "The optimal projection equations for fixed-order dynamic compensation," *IEEE Trans. Auto. Contr.*, AC-29, 1034-1037, (1984).
15. Johnson, T. L., "Optimization of low order compensators for infinite dimensional systems," *Proc. 9th IFIP Symposium on Optimization Techniques*, Warsaw, Poland, September, 1979, 394-401.
16. Pearson, R. K., "Optimal fixed form compensators for large space structures," in *ACOSS Six* (Active Control of Space Structures), RADC-TR-81-289, Rome Air Development Center, Griffiss Air Force Base, New York, 1981.

17. Johnson, T. L., "Progress in modelling and control of flexible spacecraft," *J. Franklin Inst.*, 315, 495-520 (1983).
18. Bernstein, D. S. and D. C. Hyland, "The optimal projection equations for finite-dimensional fixed-order dynamic compensation of infinite dimensional systems," *SIAM J. Control and Opt.*, 24, 122-151 (1986).
19. Richter, S., "A homotopy algorithm for solving the optimal projection equations for fixed order dynamic compensation of infinite dimensional systems," *SIAM J. Control and Opt.*, 24, 122-151 (1986).
20. Greeley, S. W. and D. C. Hyland, "Reduced order compensation: LQG reduction versus optimal projection using a homotopic continuation method," *Proc. IEEE Conference on Decision and Control*, Los Angeles, CA, Dec. 9-11, 1987, 742-747.
21. Curtain, R. F. and A. J. Pritchard, *Infinite Dimensional Linear Systems Theory*, Springer-Verlag, New York, 1978.
22. Balakrishnan, A. V., *Applied Functional Analysis*, Springer-Verlag, New York, 1981.
23. Pazy, A., *Semigroups of Linear Operators and Applications to Partial Differential Equations*, Springer-Verlag, New York, 1983.
24. Smith, M. C., "On minimal order stabilization of single loop plants," *Sys. Contr. Letters*, 7, 39-40 (1986).
25. Swartz, B. K. and R. S. Varga, "Error bounds for spline and L-spline interpolation," *J. Approx. Theory*, 6, 6-49 (1972).
26. Ito, K. and F. Kappel, "A uniformly differentiable approximation scheme for delay systems using splines," Technical Report 87-94, Institute fur Mathematik, Technische Universitat Graz and Universitat Graz, Graz, Austria, 1987, and *Appl. Math. Optim.*, submitted.
27. Ito, K., "On the approximation of eigenvalues associated with functional differential equations," *J. Diff. Equ.*, 60, 285-300 (1985).

APPENDIX D: Decentralized Control

- 76. D. S. Bernstein, "Sequential Design of Decentralized Dynamic Compensators Using the Optimal Projection Equations," Int. J. Contr., Vol. 46, pp. 1569-1577, 1987.
- 121. D. S. Bernstein and W. M. Haddad, "Robust Decentralized Static Output Feedback," submitted to Sys. Contr. Lett.

Sequential design of decentralized dynamic compensators using the optimal projection equations

DENNIS S. BERNSTEIN†

The optimal projection equations for quadratically optimal centralized fixed-order dynamic compensation are generalized to the case in which the dynamic compensator has, in addition, a fixed decentralized structure. Under a stabilizability assumption for the particular feedback configuration, the resulting optimality conditions explicitly characterize each subcontroller in terms of the plant and remaining subcontrollers. This characterization associates an oblique projection with each subcontroller and suggests an iterative sequential design algorithm. The results are applied to an interconnected flexible beam example.

1. Introduction

The purpose of this note is to consider the problem of designing decentralized dynamic feedback controllers using recently obtained results on quadratically optimal fixed-order dynamic compensation (Hyland and Bernstein 1984). As in Bernussou and Titli (1982), Looze *et al.* (1978), and Singh (1981), the overall approach is to fix the structure (information pattern and order) of the linear controller and optimize the steady-state regulation cost with respect to the controller parameters. The underlying philosophy is that the ability to carry out such an optimization procedure permits the evaluation of a particular decentralized configuration which may be dictated by implementation constraints. If there is some flexibility in designing the decentralized architecture, then these results can be used to evaluate the optimal performance of each permissible configuration, and hence to determine preferable structures. Since the present paper is confined to the question of optimal regulation, trade-offs with regard to robustness in the presence of plant variations are not considered. Such trade-offs can be included, however, by utilizing the Stratonovich multiplicative white noise approach developed by Bernstein and Hyland (1985).

To further motivate our approach, consider the problem of controlling an n th-order plant \mathcal{P} by means of a decentralized dynamic compensator consisting of subcontrollers \mathcal{C}_1 and \mathcal{C}_2 . A straightforward design technique that immediately comes to mind is that of *sequential optimization* (Davison and Gesing 1979, Jamshidi 1983). To begin, ignore \mathcal{C}_2 and design \mathcal{C}_1 as a centralized controller for \mathcal{P} . Next, regard the closed-loop system consisting of \mathcal{P} and \mathcal{C}_1 as an augmented system \mathcal{P}' and design \mathcal{C}_2 as a centralized controller for \mathcal{P}' . Now *redesign* \mathcal{C}_1 to be a centralized controller for the augmented closed-loop system composed of \mathcal{P} and \mathcal{C}_2 , and so forth. One difficulty with this scheme, however, is that of *dimension*. If, for example, one were to employ LQG at each step of this algorithm, then on the first iteration \mathcal{C}_1 would have dimension n and thus \mathcal{C}_2 would have dimension $2n$. On the second iteration, \mathcal{C}_1 would require dimension $3n$ and \mathcal{C}_2 would have order $4n$, and so forth. Such

Received 15 December 1986.

† Harris Corporation, Government Aerospace Systems Division, P.O. Box 94000, Melbourne, Florida 32902, U.S.A.

difficulties can be avoided by setting $n = 0$, which essentially corresponds to static output feedback. Although easier to implement, static output feedback lacks filtering abilities such as are inherent in LQG controllers, which are purely dynamic (i.e. strictly proper).

As discussed by Sandell *et al.* (1978), p. 119, the explanation for this difficulty is provided by the 'second-guessing' phenomenon: when LQG is used, each subcontroller must consist of linear feedback, not only of estimates of the plant states but also of *estimates of the other subcontrollers' estimates*. Hence the 'optimal' controller is given by an irrational transfer function, i.e. a distributed parameter (infinite-dimensional) system. Such controllers, of course, must be ruled out since their design and implementation (except in special cases) violate physical realizability (see, for example, Bernstein and Hyland 1986).

Having thus ruled out zeroth-order and infinite-order decentralized controllers, we focus on the problem of designing purely dynamic decentralized compensators. Moreover, by invoking the constraint of fixed subcontroller order, we overcome the second-guessing phenomenon. Utilizing the parameter optimization approach thus leads to a generalization of the result obtained by Hyland and Bernstein (1984) for centralized control. In brief, it was shown in Hyland and Bernstein (1984) that the unwieldy first-order necessary conditions for fixed-order dynamic compensation can be simplified by exploiting the presence of a previously unrecognized *oblique* projection. The resulting *optimal projection equations*, which consist of a pair of modified Riccati equations and a pair of modified Lyapunov equations coupled by the optimal projection, yield insight into the structure of the optimal dynamic compensator and emphasize the breakdown of the separation principle for reduced-order controller design. For example, the optimal compensator is the projection of a full-order dynamic controller which is generally different from the LQG design. Furthermore, this full-order controller and the oblique projection are intricately related since they are simultaneously determined by the coupled design equations. An immediate consequence is the observation that stepwise schemes employing either model reduction followed by LQG or LQG followed by model reduction are generally suboptimal. For computational purposes, the optimal projection equations permit the development of novel numerical methods which operate through successive iteration of the oblique projection. (Hyland and Bernstein 1985). Such algorithms are thus philosophically and operationally distinct from gradient search methods.

The generalization of the optimal projection equations to the decentralized case is straightforward and immediate. In the optimization process each subcontroller is viewed as a centralized controller for an augmented 'plant' consisting of the *actual* plant and all other subcontrollers. It need only be observed that the necessary conditions for optimality for the decentralized problem must consist of the collection of necessary conditions obtained by optimizing over each subcontroller separately while keeping the other subcontrollers fixed. More precisely, this statement corresponds to the fact that setting the Fréchet derivative to zero is equivalent to setting the individual partial derivatives to zero. Hence it is not surprising that the optimal projection equations for the decentralized problem involve multiple oblique projections, one associated with each subcontroller. Furthermore, each subcontroller incorporates an internal model (in the sense of an oblique projection of full-order dynamics) not only of the plant but also of all other subcontrollers. The structure of the equations suggests a sequential design algorithm such as that proposed in this work.

The simplicity with which this result is obtained should not belie its relevance to the decentralized control problem. Specifically, our approach is distinct from subsystem-decomposition techniques (Ikeda and Siljak 1980, 1981, Ikeda *et al.* 1981, 1984, Lindner 1985, Linnemann 1984, Ozguner 1979, Ramakrishna and Viswanadham 1982, Saeks 1979, Sezer and Huseyin 1984, Silkak 1978, 1983) and model-reduction methods since the optimal projection equations retain the full, interconnected plant at all times. For the proposed algorithm, decomposition techniques which exploit subsystem-interconnection data can play a role by providing a starting point for subsequent iterative refinement and optimization. Decomposition methods may also play a role when very high dimensionality precludes direct solution of the optimal projection equations. These are areas for future research.

With regard to the role of the oblique projection, it should be noted that such transformations do not, in general, preserve plant characteristics such as poles, zeros, subspaces, etc. Indeed, since the oblique projection arises as a consequence of optimality, approaches that seek to retain system invariants (e.g. Uskokovic and Medanic 1985) are generally suboptimal. In addition, the complex coupling among the plant and subcontrollers via multiple oblique projections provides an additional measure for evaluating the suboptimality of the methods proposed.

The plan of the paper is as follows. The fixed-structure decentralized dynamic-compensation problem is stated in § 2 along with the generalization of the optimal projection equations. In § 3 we propose a sequential design algorithm for solving these equations and state conditions under which convergence is guaranteed. Finally, in § 4 the algorithm is applied to the 8th-order model of a pair of simply supported beams connected by a spring. For this example, we obtain a two-channel decentralized design which is 4th-order in each channel and compare its performance with the (8th-order) centralized LQG design.

2. Problem statement and main theorem

Given the controlled system

$$\dot{x}(t) = Ax(t) + \sum_{i=1}^p B_i u_i(t) + w_0(t) \quad (2.1)$$

$$y_i(t) = C_i x(t) + w_i(t), \quad i = 1, \dots, p \quad (2.2)$$

design a fixed-structure decentralized dynamic compensator

$$\dot{x}_{ci}(t) = A_{ci} x_{ci}(t) + B_{ci} y_i(t), \quad i = 1, \dots, p \quad (2.3)$$

$$u_i(t) = C_{ci} x_{ci}(t), \quad i = 1, \dots, p \quad (2.4)$$

which minimizes the steady-state performance criterion

$$J(A_{c1}, B_{c1}, C_{c1}, \dots, A_{cp}, B_{cp}, C_{cp}) \triangleq \lim_{t \rightarrow \infty} \mathbb{E} \left[x(t)^T R_0 x(t) + \sum_{i=1}^p u_i(t)^T R_i u_i(t) \right] \quad (2.5)$$

where, for $i = 1, \dots, p$: $x \in \mathbb{R}^n$, $u_i \in \mathbb{R}^{m_i}$, $y_i \in \mathbb{R}^{l_i}$, $c_{ci} \in \mathbb{R}^{n_{ci}}$, $n_c \triangleq \sum_{i=1}^p n_{ci}$, $n_{ci} \leq n + n_c - n_{ci}$. A , B_i , C_i , A_{ci} , B_{ci} , C_{ci} , R_0 and R_i are matrices of appropriate dimension with R_0 (symmetric) non-negative definite and R_i (symmetric) positive definite; w_0 is white disturbance noise with $n \times n$ non-negative-definite intensity V_0 , and w_i is white

observation noise with $l_i \times l_i$ positive-definite intensity V_i , where w_0, w_1, \dots, w_p are mutually uncorrelated and have zero mean. \mathbb{E} denotes expectation and superscript T indicates transpose.

To guarantee that J is finite and independent of initial conditions we restrict our attention to the set of admissible stabilizing compensators

$$\mathcal{A} \triangleq \{(A_{c1}, B_{c1}, C_{c1}, \dots, A_{cp}, B_{cp}, C_{cp}) : \tilde{A} \text{ is asymptotically stable}\}$$

where the closed-loop dynamics matrix \tilde{A} is given by

$$\tilde{A} \triangleq \begin{bmatrix} A & \tilde{B}C_c \\ B_c \tilde{C} & A_c \end{bmatrix}$$

where

$$\tilde{B} \triangleq [B_1 \quad \dots \quad B_p], \quad \tilde{C} \triangleq \begin{bmatrix} C_1 \\ \vdots \\ C_p \end{bmatrix}$$

$$A_c \triangleq \text{block-diagonal}(A_{c1}, \dots, A_{cp})$$

$$B_c \triangleq \text{block-diagonal}(B_{c1}, \dots, B_{cp})$$

$$C_c \triangleq \text{block-diagonal}(C_{c1}, \dots, C_{cp})$$

(For possibly non-square matrices S_1, S_2 , block-diagonal (S_1, S_2) denotes the matrix $\begin{bmatrix} S_1 & 0 \\ 0 & S_2 \end{bmatrix}$.)

It is possible that for certain decentralized structures the system is not stabilizable, i.e. \mathcal{A} is empty (Wang and Davison 1973, Seraji 1982, Sezer and Siljak 1981). Our approach, however, is to assume that \mathcal{A} is not empty and characterize the optimal decentralized controller over the stabilizing class. Since the value of J is independent of the internal realization of each subcompensator, without loss of generality we can further restrict our attention to

$$\mathcal{A}_+ \triangleq \{(A_{c1}, B_{c1}, C_{c1}, \dots, A_{cp}, B_{cp}, C_{cp}) \in \mathcal{A} : (A_{ci}, B_{ci}) \text{ is controllable and } (C_{ci}, A_{ci}) \text{ is observable, } i = 1, \dots, p\}$$

The following lemma is an immediate consequence of Theorem 6.2.5, p. 123 of Rao and Mitra (1971). Let I_r denote the $r \times r$ identity matrix.

Lemma 2.1

Suppose $\hat{Q}, \hat{P} \in \mathbb{R}^{q \times q}$ are non-negative definite and $\text{rank } \hat{Q}\hat{P} = r$. Then there exist $G, \Gamma \in \mathbb{R}^{r \times q}$ and invertible $M \in \mathbb{R}^{r \times r}$ such that

$$\hat{Q}\hat{P} = G^T M \Gamma \quad (2.6)$$

$$\Gamma G^T = I_r \quad (2.7)$$

For convenience in stating the main theorem, call (G, M, Γ) satisfying (2.6), (2.7) a

projective factorization of $\hat{Q}\hat{P}$. Such a factorization is unique modulo an arbitrary change in basis in \mathbb{R}^r , which corresponds to nothing more than a change of basis for the internal representation of the compensator (or subcompensators in the present context).

We shall also require the following notation. Let \tilde{A}_i denote \tilde{A} with the rows and columns containing A_{ci} deleted. Similarly, let \tilde{R}_i be obtained by deleting the rows and columns corresponding to $C_{ci}^T R_i C_{ci}$ in the matrix

$$\tilde{R} \triangleq \text{block-diagonal} (R_0, C_{c1}^T R_1 C_{c1}, \dots, C_{cp}^T R_p C_{cp})$$

And furthermore, \tilde{V}_i is obtained by deleting the rows and columns containing $B_{ci} V_i B_{ci}^T$ in

$$\tilde{V} \triangleq \text{block-diagonal} (V_0, B_{c1} V_1 B_{c1}^T, \dots, B_{cp} V_p B_{cp}^T)$$

Also define

$$\tilde{B}_i \triangleq \begin{bmatrix} B_i \\ 0_{(n_c - n_{ci}) \times m_i} \end{bmatrix}, \quad \tilde{C}_i \triangleq [C_i \quad 0_{l_i \times (n_c - n_{ci})}]$$

where $0_{r \times s}$ denotes the $r \times s$ zero matrix. Note that \tilde{A}_i , \tilde{B}_i , \tilde{C}_i , \tilde{R}_i and \tilde{V}_i essentially represent the closed-loop system minus the i th subcontroller as controlled by the latter. Finally, define

$$\Sigma_i \triangleq \tilde{B}_i R_i^{-1} \tilde{B}_i^T, \quad \tilde{\Sigma}_i \triangleq \tilde{C}_i^T V_i^{-1} \tilde{C}_i$$

and, for $\tau \in \mathbb{R}^{r \times r}$, let

$$\tau_{\perp} \triangleq I_r - \tau$$

Main theorem

Suppose $(A_{c1}, B_{c1}, C_{c1}, \dots, A_{cp}, B_{cp}, C_{cp}) \in \mathcal{A}_+$ solves the steady-state fixed-structure decentralized dynamic-compensation problem. Then for $i = 1, \dots, p$ there exist $(n + n_c - n_{ci}) \times (n + n_c - n_{ci})$ non-negative-definite matrices Q_i , P_i , \hat{Q}_i and \hat{P}_i such that A_{ci} , B_{ci} and C_{ci} are given by

$$A_{ci} = \Gamma_i (\tilde{A}_i - Q_i \tilde{\Sigma}_i - \Sigma_i P_i) G_i^T \quad (2.8)$$

$$B_{ci} = \Gamma_i Q_i \tilde{C}_i^T V_i^{-1} \quad (2.9)$$

$$C_{ci} = -R_i^{-1} \tilde{B}_i^T P_i G_i^T \quad (2.10)$$

for some projective factorization G_i , M_i , Γ_i of $\hat{Q}_i \hat{P}_i$, and such that, with $\tau_i = G_i^T \Gamma_i$, the following conditions are satisfied:

$$0 = \tilde{A}_i Q_i + Q_i \tilde{A}_i^T + \tilde{V}_i - Q_i \tilde{\Sigma}_i Q_i + \tau_{i\perp} Q_i \tilde{\Sigma}_i Q_i \tau_{i\perp}^T \quad (2.11)$$

$$0 = \tilde{A}_i^T P_i + P_i \tilde{A}_i + \tilde{R}_i - P_i \tilde{\Sigma}_i P_i + \tau_{i\perp}^T P_i \tilde{\Sigma}_i P_i \tau_{i\perp} \quad (2.12)$$

$$0 = (\tilde{A}_i - \Sigma_i P_i) \hat{Q}_i + \hat{Q}_i (\tilde{A}_i - \Sigma_i P_i)^T + Q_i \tilde{\Sigma}_i Q_i - \tau_{i\perp} Q_i \tilde{\Sigma}_i Q_i \tau_{i\perp}^T \quad (2.13)$$

$$0 = (\tilde{A}_i - Q_i \tilde{\Sigma}_i)^T \hat{P}_i + \hat{P}_i (\tilde{A}_i - Q_i \tilde{\Sigma}_i) + P_i \tilde{\Sigma}_i P_i - \tau_{i\perp}^T P_i \tilde{\Sigma}_i P_i \tau_{i\perp} \quad (2.14)$$

$$\text{rank } \hat{Q}_i = \text{rank } \hat{P}_i = \text{rank } \hat{Q}_i \hat{P}_i = n_{ci} \quad (2.15)$$

Remark 2.1

Because of (2.7) the matrix τ_i is idempotent, i.e. $\tau_i^2 = \tau_i$. This projection corresponding to the i th subcontroller is an *oblique* projection (as opposed to an orthogonal projection) since it is not necessarily symmetric. Furthermore, τ_i is given in closed form by

$$\tau_i = \hat{Q}_i \hat{P}_i (\hat{Q}_i \hat{P}_i)^*$$

where $()^*$ denotes the (Drazin) group generalized inverse (see, for example, Campbell and Meyer, 1979, p. 124).

3. Proposed algorithm*Sequential design algorithm*

- Step 1.** Choose a starting point consisting of initial subcontroller designs;
- Step 2.** For a sequence $\{i_k\}_{k=1}^\infty$, where $i_k \in \{1, \dots, p\}$, $k = 1, 2, \dots$, redesign subcontroller i_k as an optimal fixed-order centralized controller for the plant and remaining subcontrollers;
- Step 3.** Compute the cost J_k of the current design and check $J_k - J_{k-1}$ for convergence.

Note that the first two steps of the algorithm consist of (i) bringing suboptimal subcontrollers 'on line' and (ii) iteratively refining each subcontroller. As discussed in § 1, the choice of a starting design for Step 1 can be obtained by a variety of existing methods such as subsystem decomposition. As for subcontroller refinement, note that each subcontroller redesign procedure is equivalent to replacing a suboptimal subcontroller with a subcontroller which is optimal with respect to the plant and remaining subcontrollers.

Proposition 3.1

For a given starting design and redesign sequence $\{i_k\}_{k=1}^\infty$ suppose that the optimal projection equations can be solved for each k to yield the global minimum. Then $\{J_k\}_{k=1}^\infty$ is monotonically non-increasing and hence convergent.

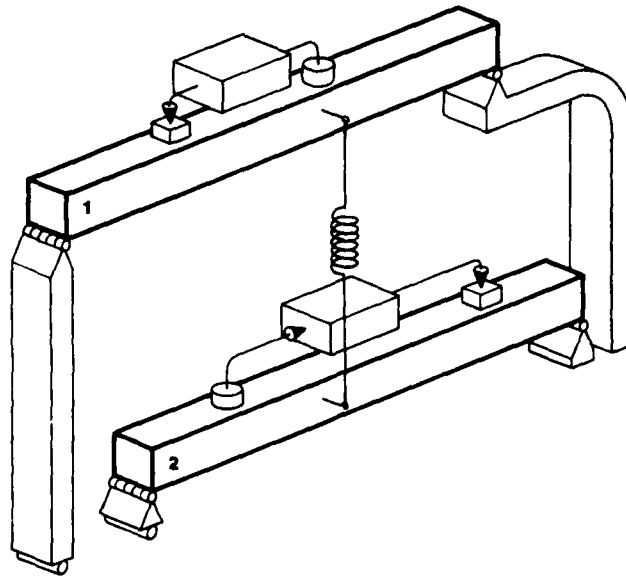
Determining both a suitable starting point and redesign sequence for solvability and attaining the decentralized global minimum remain areas for future research. With regard to algorithms for solving the optimal projection equations for each subcontroller redesign procedure, details of proposed algorithms can be found in the works of Hyland (1983, 1984) and Hyland and Bernstein (1985).

4. Application to interconnected flexible beams

To demonstrate the applicability of the main theorem and the sequential design algorithm, we consider a pair of simply supported Euler-Bernoulli flexible beams interconnected by a spring (see the Figure). Each beam possesses one rate sensor and one force actuator. Retaining two vibrational modes in each beam, we obtain the 8th-order interconnected model

$$A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}, \quad B_1 = \begin{bmatrix} B_{11} \\ 0_{4 \times 1} \end{bmatrix}, \quad B_2 = \begin{bmatrix} 0_{4 \times 1} \\ B_{22} \end{bmatrix}$$

$$C_1 = [C_{11} \quad 0_{1 \times 4}], \quad C_2 = [0_{1 \times 4} \quad C_{22}]$$



where

$$A_{ii} = \begin{bmatrix} 0 & \omega_{1i} & 0 & 0 \\ -\omega_{1i} - (k/\omega_{1i})(\sin \pi c_i)^2 & -2\zeta_i \omega_{1i} & -(k/\omega_{2i})(\sin \pi c_i)(\sin 2\pi c_i) & 0 \\ 0 & 0 & 0 & \omega_{2i} \\ -(k/\omega_{1i})(\sin \pi c_i)(\sin 2\pi c_i) & 0 & -\omega_{2i} - (k/\omega_{2i})(\sin 2\pi c_i)^2 & -2\zeta_i \omega_{2i} \end{bmatrix}$$

$$A_{ij} = \begin{bmatrix} 0 & 0 & 0 & 0 \\ (k/\omega_{1j})(\sin \pi c_i)(\sin \pi c_j) & 0 & (k/\omega_{2j})(\sin \pi c_i)(\sin 2\pi c_j) & 0 \\ 0 & 0 & 0 & 0 \\ (k/\omega_{1j})(\sin \pi c_j)(\sin 2\pi c_i) & 0 & (k/\omega_{2j})(\sin 2\pi c_i)(\sin 2\pi c_j) & 0 \end{bmatrix}$$

$i \neq j$

$$B_{ii} = \begin{bmatrix} 0 \\ -\sin \pi a_i \\ 0 \\ -\sin 2\pi a_i \end{bmatrix}, \quad C_{ii} = [0 \quad \sin \pi s_i \quad 0 \quad \sin 2\pi s_i]$$

$$a_i = \hat{a}_i/L_i, \quad s_i = \hat{s}_i/L_i, \quad c_i = \hat{c}_i/L_i$$

In the above definitions, k is the spring constant, ω_{ji} is the j th modal frequency of the i th beam, ζ_i is the damping ratio of the i th beam, L_i is the length of the i th beam, and \hat{a}_i , \hat{s}_i and \hat{c}_i are, respectively, the actuator, sensor and spring-connection coordinates as measured from the left in the Figure. The chosen values are

$$k = 10$$

$$\omega_{1i} = 1, \quad \omega_{2i} = 4, \quad \zeta_i = 0.005, \quad L_i = 1, \quad i = 1, 2$$

$$\hat{a}_1 = 0.3, \quad \hat{s}_1 = 0.65, \quad \hat{c}_1 = 0.6$$

$$\hat{a}_2 = 0.8, \quad \hat{s}_2 = 0.2, \quad \hat{c}_2 = 0.4$$

In addition, weighting and intensity matrices are chosen to be

$$R_1 = \text{block-diagonal} \left(\begin{bmatrix} 1 & 0 \\ 0 & 1/\omega_{11} \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 0 & 1/\omega_{21} \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 0 & 1/\omega_{12} \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 0 & 1/\omega_{22} \end{bmatrix} \right)$$

$$R_2 = R_3 = 0.1I_2$$

$$V_0 = \text{block-diagonal} \left(\begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \right)$$

$$V_1 = V_2 = 0.1I_2$$

For this problem the open-loop cost was evaluated and the centralized 8th-order LQG design was obtained to provide a baseline. To provide a starting point for the sequential design algorithm, a pair of 4th-order LQG controllers were designed for each beam separately ignoring the interconnection, i.e. setting $k = 0$. The optimal projection equations were then utilized to iteratively refine each subcontroller. The results are summarized in the Table.

| Design | Cost |
|--------------------------|-------|
| Open loop | 163.5 |
| Centralized LQG | |
| $n_c = 8$ | 19.99 |
| Suboptimal decentralized | |
| $n_{c1} = n_{c2} = 4$ | 59.43 |
| Redesign subcontroller 2 | 28.19 |
| Redesign subcontroller 1 | 23.29 |
| Redesign subcontroller 2 | 23.04 |
| Redesign subcontroller 1 | 22.25 |
| Redesign subcontroller 2 | 21.94 |
| Redesign subcontroller 1 | 21.86 |
| Redesign subcontroller 2 | 21.81 |
| Redesign subcontroller 1 | 21.79 |

ACKNOWLEDGMENTS

This research was supported in part by the Air Force Office of Scientific Research under contracts F49620-86-C-0002 and F49620-86-C-0038.

The author wishes to thank Scott W. Greeley for providing the beam model in § 4 and for carrying out the design computations.

REFERENCES

- BERNSTEIN, D. S., and HYLAND, D. C., 1985, *Proc. 24th I.E.E.E. Conf. Dec. Control*, 745; 1986, *SIAM JI Control Optim.*, **24**, 122.
- BERNUSSOU, J., and TITLI, A., 1982, *Interconnected Dynamical Systems: Stability, Decomposition and Decentralization* (New York: North-Holland).
- CAMPBELL, S. L., and MEYER, C. D., Jr., 1979, *Generalized Inverses of Linear Transformations* (London: Pitman).
- DAVISON, E. J., and GESING, W., 1979, *Automatica*, **15**, 307.
- DE CARLO, R. A., and SAEKS, R., 1981, *Interconnected Dynamical Systems* (New York: Marcel Dekker).
- HYLAND, D. C., 1983, *AIAA 21st Aerospace Sciences Meeting*, paper 83-0303; 1984, *Proc. AIAA Dynamics Specialists Conf.*, 381.
- HYLAND, D. C., and BERNSTEIN, D. S., 1984, *I.E.E.E. Trans. autom. Control*, **29**, 1034; 1985, *Ibid.*, **30**, 1201.
- IKEDA, M., and SILJAK, D. D., 1980, *Large-scale Syst.*, **1**, 29; 1981, *I.E.E.E. Trans. autom. Control*, **26**, 1118.
- IKEDA, M., SILJAK, D. D., and WHITE, D. E., 1981, *J. optim. Theory Applic.*, **34**, 279; 1984, *I.E.E.E. Trans. autom. Control*, **29**, 244.
- JAMSHIDI, M., 1983, *Large-Scale Systems* (New York: North-Holland).
- LINDNER, D. K., 1985, *Syst. Control Lett.*, **6**, 109.
- LINNEMANN, A., 1984, *I.E.E.E. Trans. autom. Control*, **29**, 1052.
- LOOZE, D. P., HOUP, P. K., SANDELL, N. R., and ATHANS, M., 1978, *I.E.E.E. Trans. autom. Control*, **23**, 268.
- OZGUNER, U., 1979, *I.E.E.E. Trans. autom. Control*, **24**, 652.
- RAMAKRISHNA, A., and VISWANADHAM, N., 1982, *I.E.E.E. Trans. autom. Control*, **27**, 159.
- RAO, C. R., and MITRA, S. K., 1971, *Generalized Inverse of Matrices and its Applications* (New York: Wiley).
- SAEKS, R., 1979, *I.E.E.E. Trans. autom. Control*, **24**, 269.
- SANDELL, N., VARAIYA, P., ATHANS, M., and SAFONOV, M. G., 1978, *I.E.E.E. Trans. autom. Control*, **23**, 108.
- SERAJI, H., 1982, *Int. J. Control*, **35**, 775.
- SEZER, M. E., and HUSEYIN, O., 1984, *Automatica*, **16**, 205.
- SEZER, M. E., and SILJAK, D. D., 1981, *Syst. Control Lett.*, **1**, 60.
- SILJAK, D. D., 1978, *Large Scale Dynamic Systems* (New York: North-Holland); 1983, *Large-Scale Syst.*, **4**, 279.
- SINGH, M. G., 1981, *Decentralised Control* (New York: North-Holland).
- USKOKOVIC, Z., and MEDANIC, J., 1985, *Proc. 24th I.E.E.E. Conf. on Decision and Control*, 837.
- WANG, S. H., and DAVISON, E. J., 1973, *I.E.E.E. Trans. autom. Control*, **18**, 473.

April 1988

**Robust Decentralized Output Feedback:
The Static Controller Case**

by

Dennis S. Bernstein
Harris Corporation
Government Aerospace Systems Division
MS 22/4848
Melbourne, FL 32902

Wassim M. Haddad
Department of Mechanical and
Aerospace Engineering
Florida Institute of Technology
Melbourne, FL 32901

Keywords: robust, guaranteed bounds, optimal, structured uncertainty, real parameters

Abstract

Sufficient conditions are developed for designing robust decentralized static output feedback controllers. The approach involves deriving necessary conditions for minimizing a bound on closed-loop performance over a specified range of uncertain parameters. The effect of plant parameter variations on the closed-loop covariance is overbounded by means of a modified Lyapunov equation whose solutions are guaranteed to provide robust stability and performance.

Supported in part by the Air Force Office of Scientific Research under contract F49620-86-C-0038.

1. Introduction

Because of implementation constraints, cost, and reliability considerations, a decentralized controller architecture is often required for controlling large scale systems. Furthermore, such controllers must be robust to variations in plant parameters. The present paper addresses both of these concerns within the context of a robust decentralized theory for continuous-time static controllers.

The approach to controller design considered herein involves optimizing closed-loop performance with respect to the feedback gains. This approach to output feedback was studied for centralized controllers in [8,9] and for decentralized controllers in [10]. An interesting feature of [9,10] is the recognition of an oblique projection (idempotent matrix) which allows the necessary conditions to be written in terms of a modified Riccati equation. When the problem is specialized to full-state feedback, the projection becomes the identity and the modified Riccati equation coincides with the standard Riccati equation of LQR theory. It should be pointed out that this oblique projection is distinct from the oblique projection arising in dynamic compensation ([7]). A unified treatment of the static/dynamic (nonstrictly proper) centralized control problem involving both projections is given in [2].

The present paper goes beyond earlier work by deriving sufficient conditions for robust stability and performance with respect to variations in the plant parameters. Although plant disturbances are represented in the usual stochastic manner by means of additive white noise, uncertainty in the plant dynamics is modeled deterministically by means of constant structured parameter variations within bounded sets. Thus, for example, the dynamics matrix A is replaced by $A + \sum_{k=1}^p \sigma_k A_k$, where σ_k is a constant uncertain parameter assumed only to lie within the interval $[-\alpha_k, \alpha_k]$ but otherwise unknown, and A_k is a fixed matrix denoting the structure of the uncertain parameter σ_k as it appears in the nominal dynamics matrix A . The system performance is defined to be the worst-case value over the class of parameter uncertainties of a quadratic criterion averaged over the disturbance statistics.

Since the closed-loop performance can be written in terms of the second-moment matrix, a performance bound over the class of uncertain parameters can be obtained by bounding the state covariance. The key to bounding the state covariance is to replace the usual Lyapunov equation for the second-moment matrix by a modified Lyapunov equation. In the present paper the modified Lyapunov equation is constructed by adding two additional terms. The first term corresponds to a

uniform right shift of the open-loop dynamics. As is well known ([1]), such a shift may arise from an exponential performance weighting and leads to a uniform stability margin for the closed-loop system. In order to guarantee robustness with respect to specified structured parameter variations, however, an additional term of the form $A_k Q A_k^T$ is required. Such terms arise naturally in systems with multiplicative white noise; see [3,4] and the references therein for further details. The exponential cost weighting and multiplicative noise interpretations for the uncertainty bound have no bearing in the present paper, however, since parameter variations are modeled deterministically as constant variations within bounded sets.

Having bounded the state covariance over the class of parameter uncertainties, the worst-case performance can thus be bounded in terms of the solution of the modified Lyapunov equation. The performance bound can be viewed as an *auxiliary cost* and thus leads to the Auxiliary Minimization Problem: Minimize the performance bound while satisfying the modified Lyapunov equation. The nice feature of the auxiliary problem is that necessary conditions for optimality of the performance bound now serve as *sufficient* conditions for robust performance in the original problem. Thus our approach seeks to rectify one of the principal drawbacks of necessity theory, namely, guarantees of stability and performance. Furthermore, it should be noted that if numerical solution of the optimality conditions yields a local extremal which is not the global optimum, then robust stability and performance are still guaranteed, although the performance of the extremal may not be as good as the performance provided by the global minimum. Philosophically, the overall approach of control design for a performance bound is related to guaranteed cost control ([6]). We note, however, that the bound utilized in [6] is nondifferentiable, which precludes the approach of the present paper.

A further extension of previous approaches considered in the present paper involves the types of feedback loops considered. Specifically, the usual approach to static output feedback involves nonnoisy measurements and weighted controls, while the dual problem involves feeding back noisy measurements to unweighted controls. This situation leads to an additional projection ([2]) which is dual to the projection discussed in [9,10]. The inclusion of the dual case now leads to a *pair* of modified Riccati equations coupled by both the uncertainty bounds and the oblique projections.

In addition to the two types of loops discussed above, one may wish to consider the two remaining cases, namely, feeding back noisy measurements to weighted controls and feeding back nonnoisy measurements to unweighted controls. It is easy to show, however, that the former case

leads to an undefined (i.e., infinite) value for the performance while the latter case is highly singular and thus will not be treated here.

Finally, the scope of the present paper is limited to the development of sufficient conditions for robust decentralized output feedback. Numerical solution of these equations can be carried out by extending available algorithms for centralized output feedback. Numerical algorithms for solving a single modified Riccati equation in the absence of uncertainty bounds are discussed in [10].

2. Notation and Definitions

| | |
|---|--|
| $\mathbb{R}, \mathbb{R}^{r \times s}, \mathbb{R}^r, \mathbb{E}$ | real numbers, $r \times s$ real numbers, $\mathbb{R}^{r \times 1}$, expectation |
| $I_r, (\)^T$ | $r \times r$ identity, transpose |
| \oplus, \otimes | Kronecker sum, Kronecker product ([5]) |
| \mathbb{S}^r | $r \times r$ symmetric matrices |
| \mathbb{N}^r | $r \times r$ symmetric nonnegative-definite matrices |
| \mathbb{P}^r | $r \times r$ symmetric positive-definite matrices |
| $Z_1 \leq Z_2$ | $Z_2 - Z_1 \in \mathbb{N}^r, \quad Z_1, Z_2 \in \mathbb{S}^r$ |
| $Z_1 < Z_2$ | $Z_2 - Z_1 \in \mathbb{P}^r, \quad Z_1, Z_2 \in \mathbb{S}^r$ |
| asymptotically stable matrix | matrix with eigenvalues in open left half plane |
| n, r, s, p | positive integers |
| i, j, k | indices, $i = 1, \dots, r, \quad j = 1, \dots, s, \quad k = 1, \dots, p$ |
| $m_i, \hat{\ell}_i$ | positive integers, $i = 1, \dots, r$ |
| \hat{m}_j, ℓ_j | positive integers, $j = 1, \dots, s$ |
| x | n -dimensional vector |
| u_i, \hat{y}_i | $m_i, \hat{\ell}_i$ -dimensional vectors, $i = 1, \dots, r$ |
| \hat{u}_j, y_j | \hat{m}_j, ℓ_j -dimensional vectors, $j = 1, \dots, s$ |
| $A, \Delta A$ | $n \times n$ matrices |
| $B_i, \Delta B_i; \hat{C}_i$ | $n \times m_i$ matrices; $\hat{\ell}_i \times n$ matrices, $i = 1, \dots, r$ |
| $\hat{B}_j; C_j, \Delta C_j$ | $n \times \hat{m}_j$ matrices; $\ell_j \times n$ matrices, $j = 1, \dots, s$ |
| A_k | $n \times n$ matrices, $k = 1, \dots, p$ |
| B_{ik} | $n \times m_i$ matrices, $i = 1, \dots, r, \quad k = 1, \dots, p$ |
| C_{jk} | $\ell_j \times n$ matrices, $j = 1, \dots, s, \quad k = 1, \dots, p$ |
| D_{ci} | $m_i \times \hat{\ell}_i$ matrices, $i = 1, \dots, r$ |
| E_{cj} | $\hat{m}_j \times \ell_j$ matrices, $j = 1, \dots, s$ |
| α | positive number |
| A_α | $A + \frac{\alpha}{2} I_n$ |
| α_k | positive number, $k = 1, \dots, p$ |
| γ_k | $\alpha_k^2 / \alpha, \quad k = 1, \dots, p$ |
| σ_k | real number, $k = 1, \dots, p$ |
| $w_0(t), w_j(t)$ | n -dimensional, ℓ_j -dimensional white noise, $j = 1, \dots, s$ |

| | |
|-------------------------------|--|
| V_0, V_j | intensities of w_0, w_j ; $V_0 \in \mathbb{R}^n$, $V_j \in \mathbb{R}^{\ell_j}$, $j = 1, \dots, s$ |
| V_{0j} | $n \times \ell_j$ cross intensity of w_0, w_j , $j = 1, \dots, s$ |
| R_0, R_i | state and control weightings; $R_0 \in \mathbb{R}^n$, $R_i \in \mathbb{R}^{m_i}$, $i = 1, \dots, r$ |
| R_{0i} | $n \times m_i$ cross weighting; $R_0 - R_{0i} R_i^{-1} R_{0i}^T \geq 0$, $i = 1, \dots, r$ |
| $\tilde{A}, \tilde{A}_\alpha$ | $A + \sum_{i=1}^r B_i D_{\alpha i} \hat{C}_i + \sum_{j=1}^s \hat{B}_j E_{cj} C_j$ $\tilde{A} + \frac{\alpha}{2} I_n$ |
| $\Delta \tilde{A}$ | $\Delta A + \sum_{i=1}^r \Delta B_i D_{\alpha i} \hat{C}_i + \sum_{j=1}^s \hat{B}_j E_{cj} \Delta C_j$ |
| $\tilde{w}(t)$ | $w_0(t) + \sum_{j=1}^s \hat{B}_j E_{cj} w_j(t)$ |
| \tilde{R} | $R_0 + \sum_{i=1}^r [R_{0i} D_{\alpha i} \hat{C}_i + \hat{C}_i^T D_{\alpha i}^T R_{0i}^T + \hat{C}_i^T D_{\alpha i}^T R_i D_{\alpha i} \hat{C}_i]$ |
| \tilde{V} | $V_0 + \sum_{j=1}^s [V_{0j} E_{cj}^T \hat{B}_j^T + \hat{B}_j E_{cj} V_{0j}^T + \hat{B}_j E_{cj} V_j E_{cj}^T \hat{B}_j^T]$ |

For arbitrary $n \times n$ Q, P define:

$$R_{\alpha i} \triangleq R_i + \sum_{k=1}^p \gamma_k B_{ik}^T P B_{ik}, \quad P_{\alpha i} \triangleq B_i^T P + R_{0i}^T + \sum_{k=1}^p \gamma_k B_{ik}^T P A_k, \quad i = 1, \dots, r,$$

$$V_{\alpha j} \triangleq V_j + \sum_{k=1}^p \gamma_k C_{jk} Q C_{jk}^T, \quad Q_{\alpha j} \triangleq Q C_j^T + V_{0j} + \sum_{k=1}^p \gamma_k A_k Q C_{jk}^T, \quad j = 1, \dots, s.$$

3. Robust Stability and Performance Problem

In this section we state the Robust Stability and Performance Problem along with related notation for later use. Let

$$\mathcal{U} \subset \mathbb{R}^{n \times n} \times \mathbb{R}^{n \times m_1} \times \dots \times \mathbb{R}^{n \times m_r} \times \mathbb{R}^{\ell_1 \times n} \times \dots \times \mathbb{R}^{\ell_s \times n}$$

denote the set of uncertain perturbations $(\Delta A, \Delta B_1, \dots, \Delta B_r, \Delta C_1, \dots, \Delta C_s)$ of the nominal system matrices $A, B_1, \dots, B_r, C_1, \dots, C_s$.

Robust Stability and Performance Problem. Determine $(D_{c1}, \dots, D_{cr}, E_{c1}, \dots, E_{cs})$ such that the closed-loop system consisting of the n th-order controlled and disturbed plant

$$\dot{x}(t) = (A + \Delta A)x(t) + \sum_{i=1}^r (B_i + \Delta B_i)u_i(t) + \sum_{j=1}^s \hat{B}_j \hat{u}_j(t) + w_0(t), \quad t \in [0, \infty), \quad (3.1)$$

nonnoisy and noisy measurements

$$\hat{y}_i(t) = \hat{C}_i x(t), \quad i = 1, \dots, r, \quad (3.2)$$

$$y_j(t) = (C_j + \Delta C_j)x(t) + w_j(t), \quad j = 1, \dots, s, \quad (3.3)$$

and static output feedback controller

$$u_i(t) = D_{ci} \hat{y}_i(t), \quad i = 1, \dots, r, \quad (3.4)$$

$$\hat{u}_j(t) = E_{c_j} y_j(t), \quad j = 1, \dots, s, \quad (3.5)$$

is asymptotically stable for all variations in \mathcal{U} and the performance criterion

$$J(D_{c1}, \dots, D_{cr}, E_{c1}, \dots, E_{cs}) \triangleq \sup_{\mathcal{U}} \limsup_{t \rightarrow \infty} \mathbb{E} \left[x^T(t) R_0 x(t) + 2 \sum_{i=1}^r x^T(t) R_{0i} u_i(t) + \sum_{i=1}^r u_i^T(t) R_i u_i(t) \right] \quad (3.6)$$

is minimized.

For each controller $(D_{c1}, \dots, D_{cr}, E_{c1}, \dots, E_{cs})$ and variation in \mathcal{U} , the closed-loop system (3.1)–(3.5) is given by

$$\dot{x}(t) = (\tilde{A} + \Delta \tilde{A})x(t) + \tilde{w}(t), \quad t \in [0, \infty), \quad (3.7)$$

where $\tilde{w}(t)$ is white noise with intensity $\tilde{V} \in \mathbb{R}^n$.

Remark 3.1. In the case $\Delta A, \Delta B_i, \Delta C_j = 0$ it is well known that stabilizability is related to the existence of fixed modes ([11]). When plant uncertainties are present the problem is, of course, far more complex. In the present paper sufficient conditions for robust stability are obtained as a consequence of the existence of robust performance bounds.

Remark 3.2. Note that the controller architecture is quite general in that it includes two distinctly different types of decentralized loops. The first type, indexed by $i = 1, \dots, r$, involves feeding back nonnoisy measurements to weighted controls. This is the standard setting in the optimal output-feedback literature ([8–10]). In addition, we include the dual situation, indexed by $j = 1, \dots, s$, which involves feeding back noisy measurements to unweighted controls. The case in which only one type of loop is present can be formally recovered from our results by ignoring B_i and \hat{C}_i or \hat{B}_j and C_j as required. As noted in Section 1, noisy measurements cannot be fed back to weighted controls via static control, while feeding back nonnoisy measurements to unweighted controls is a singular problem.

Remark 3.3. Note that the problem statement is restrictive in the sense that uncertainties in both the control and observation matrices are not permitted within the same feedback loop. Although it is indeed possible to permit such simultaneous uncertainties, the development is considerably more complex and hence is not treated here.

Remark 3.4. The cost functional (3.6) is identical to the LQG criterion (usually stated in terms of an averaged integral) with the exception of the supremum for evaluating worst case over \mathcal{U} .

4. Sufficient Conditions for Robust Stability and Performance

In practice, steady-state performance is only of interest when the closed-loop system (3.7) is stable over \mathcal{U} . The following result, which expresses the performance in terms of the state covariance, is immediate.

Lemma 4.1. Let $(D_{c1}, \dots, D_{cr}, E_{c1}, \dots, E_{cs})$ be given and suppose the system (3.7) is stable for all variations in \mathcal{U} . Then

$$J(D_{c1}, \dots, D_{cr}, E_{c1}, \dots, E_{cs}) = \sup_{\mathcal{U}} \text{tr } Q_{\Delta \tilde{A}} \tilde{R}, \quad (4.1)$$

where $Q_{\Delta \tilde{A}} \triangleq \lim_{t \rightarrow \infty} \mathbb{E}[x(t)x^T(t)] \in \mathbb{R}^n$ is the unique solution to

$$0 = (\tilde{A} + \Delta \tilde{A})Q_{\Delta \tilde{A}} + Q_{\Delta \tilde{A}}(\tilde{A} + \Delta \tilde{A})^T + \tilde{V}. \quad (4.2)$$

Remark 4.1. When \mathcal{U} is compact, "sup" in (4.1) can be replaced by "max".

We now seek upper bounds for $J(D_{c1}, \dots, D_{cr}, E_{c1}, \dots, E_{cs})$. Our assumptions allow us to obtain robust stability as a consequence of robust performance.

Theorem 4.1. Let $\Omega : \mathbb{R}^n \times \mathbb{R}^{m_1 \times \ell_1} \times \dots \times \mathbb{R}^{m_r \times \ell_r} \times \mathbb{R}^{m_1 \times \ell_1} \times \dots \times \mathbb{R}^{m_s \times \ell_s} \rightarrow \mathbb{S}^n$ be such that

$$\begin{aligned} \Delta \tilde{A}Q + Q\Delta \tilde{A}^T &\leq \Omega(Q, D_{c1}, \dots, D_{cr}, E_{c1}, \dots, E_{cs}), \\ (\Delta A, \Delta B_1, \dots, \Delta B_r, \Delta C_1, \dots, \Delta C_s) &\in \mathcal{U}, \\ (Q, D_{c1}, \dots, D_{cr}, E_{c1}, \dots, E_{cs}) &\in \mathbb{R}^n \times \mathbb{R}^{m_1 \times \ell_1} \times \dots \times \mathbb{R}^{m_r \times \ell_r} \times \mathbb{R}^{m_1 \times \ell_1} \times \dots \times \mathbb{R}^{m_s \times \ell_s}. \end{aligned} \quad (4.3)$$

Furthermore, for given $(D_{c1}, \dots, D_{cr}, E_{c1}, \dots, E_{cs})$ suppose there exists $Q \in \mathbb{R}^n$ satisfying

$$0 = \tilde{A}Q + Q\tilde{A}^T + \Omega(Q, D_{c1}, \dots, D_{cr}, E_{c1}, \dots, E_{cs}) + \tilde{V}. \quad (4.4)$$

Then the pair $(\tilde{A} + \Delta \tilde{A}, \tilde{V}^{\frac{1}{2}})$ is stabilizable for all variations in \mathcal{U} if and only if $\tilde{A} + \Delta \tilde{A}$ is asymptotically stable for all variations in \mathcal{U} . In this case,

$$Q_{\Delta \tilde{A}} \leq Q, \quad (4.5)$$

where $Q_{\Delta \tilde{A}}$ satisfies (4.2), and

$$J(D_{c1}, \dots, D_{cr}, E_{c1}, \dots, E_{cs}) \leq \text{tr } Q \tilde{R}. \quad (4.6)$$

Proof. For all variations in \mathcal{U} , (4.4) is equivalent to

$$0 = (\tilde{A} + \Delta\tilde{A})Q + Q(\tilde{A} + \Delta\tilde{A})^T + \Phi(Q, D_{c1}, \dots, D_{cr}, E_{c1}, \dots, E_{cs}, \Delta\tilde{A}) + \tilde{V}, \quad (4.7)$$

where

$$\Phi(Q, D_{c1}, \dots, D_{cr}, E_{c1}, \dots, E_{cs}, \Delta\tilde{A}) \triangleq \Omega(Q, D_{c1}, \dots, D_{cr}, E_{c1}, \dots, E_{cs}) - (\Delta\tilde{A}Q + Q\Delta\tilde{A}^T).$$

Note that by (4.3), $\Phi(\cdot) \geq 0$ for all variations in \mathcal{U} . If $(\tilde{A} + \Delta\tilde{A}, \tilde{V}^{\frac{1}{2}})$ is stabilizable for all variations in \mathcal{U} , it follows from Theorem 3.6 of [12] that $(\tilde{A} + \Delta\tilde{A}, [\tilde{V} + \Phi(Q, D_{c1}, \dots, D_{cr}, E_{c1}, \dots, E_{cs}, \Delta\tilde{A})]^{\frac{1}{2}})$ is stabilizable for all variations in \mathcal{U} . Hence Lemma 12.2 of [12] implies $\tilde{A} + \Delta\tilde{A}$ is asymptotically stable for all variations in \mathcal{U} . The converse is immediate. Next, subtracting (4.2) from (4.7) yields

$$0 = (\tilde{A} + \Delta\tilde{A})(Q - Q_{\Delta\tilde{A}}) + (Q - Q_{\Delta\tilde{A}})(\tilde{A} + \Delta\tilde{A})^T + \Phi(Q, D_{c1}, \dots, D_{cr}, E_{c1}, \dots, E_{cs}, \Delta\tilde{A}),$$

or, equivalently, (since $\tilde{A} + \Delta\tilde{A}$ is asymptotically stable)

$$Q - Q_{\Delta\tilde{A}} = \int_0^\infty e^{(\tilde{A} + \Delta\tilde{A})t} \Phi(Q, D_{c1}, \dots, D_{cr}, E_{c1}, \dots, E_{cs}, \Delta\tilde{A}) e^{(\tilde{A} + \Delta\tilde{A})^T t} dt \geq 0,$$

which implies (4.5). Finally, (4.5) and (4.1) yield (4.6). \square

Remark 4.2. If \tilde{V} is positive definite then the stabilizability hypothesis of Theorem 4.1 is automatically satisfied for all variations in \mathcal{U} .

5. Uncertainty Structure and the Quadratic Lyapunov Bound

The uncertainty set \mathcal{U} is assumed to be of the form

$$\begin{aligned} \mathcal{U} = \{ & (\Delta A, \Delta B_1, \dots, \Delta B_r, \Delta C_1, \dots, \Delta C_s) : \\ & \Delta A = \sum_{k=1}^p \sigma_k A_k, \quad \Delta B_i = \sum_{k=1}^p \sigma_k B_{ik}, \quad i = 1, \dots, r, \\ & \Delta C_j = \sum_{k=1}^p \sigma_k C_{jk}, \quad j = 1, \dots, s, \quad \sum_{k=1}^p \sigma_k^2 / \alpha_k^2 \leq 1 \}, \end{aligned} \quad (5.1)$$

where, for $k = 1, \dots, p$: $(A_k, B_{1k}, \dots, B_{rk}, C_{1k}, \dots, C_{sk})$ are fixed matrices denoting the structure of the parametric uncertainty; α_k is a given uncertainty bound; and σ_k is an uncertain parameter. Note that the uncertain parameters σ_k are assumed to lie in a specified ellipsoidal region in \mathbb{R}^p . The closed-loop system thus has structured uncertainty of the form

$$\Delta\tilde{A} = \sum_{k=1}^p \sigma_k \tilde{A}_k, \quad (5.2)$$

where

$$\tilde{A}_k \triangleq A_k + \sum_{i=1}^r B_{ik} D_{ci} \hat{C}_i + \sum_{j=1}^s \hat{B}_j E_{cj} C_{jk}, \quad k = 1, \dots, p. \quad (5.3)$$

To obtain explicit gain expressions for $(D_{c1}, \dots, D_{cr}, E_{c1}, \dots, E_{cs})$ we assume that, for each $k \in \{1, \dots, p\}$, at most one of the matrices $B_{1k}, \dots, B_{rk}, C_{1k}, \dots, C_{sk}$ is nonzero. Note that this assumption does *not* preclude the treatment of uncertainties in the input and output matrices. It requires only that such uncertainties be modeled as uncorrelated.

Given the structure of \mathcal{U} defined by (5.1), the bound Ω satisfying (4.3) can now be specified. In the following result Q denotes an arbitrary element of \mathbb{N}^n , *not* necessarily a solution of (4.4).

Proposition 5.1. Let α be an arbitrary positive scalar. Then the function

$$\Omega(Q, D_{c1}, \dots, D_{cr}, E_{c1}, \dots, E_{cs}) = \alpha Q + \alpha^{-1} \sum_{k=1}^p \alpha_k^2 \tilde{A}_k Q \tilde{A}_k^T \quad (5.4)$$

satisfies (4.3) with \mathcal{U} given by (5.1).

Proof. Note that

$$\begin{aligned} 0 &\leq \sum_{k=1}^p [(\alpha^{\frac{1}{2}} \sigma_k / \alpha_k) I_n - (\alpha_k / \alpha^{\frac{1}{2}}) \tilde{A}_k] Q [(\alpha^{\frac{1}{2}} \sigma_k / \alpha_k) I_n - (\alpha_k / \alpha^{\frac{1}{2}}) \tilde{A}_k]^T \\ &= \alpha \sum_{k=1}^p (\sigma_k^2 / \alpha_k^2) Q + \alpha^{-1} \sum_{k=1}^p \alpha_k^2 \tilde{A}_k Q \tilde{A}_k^T - \sum_{k=1}^p \sigma_k (\tilde{A}_k Q + Q \tilde{A}_k^T), \end{aligned}$$

which yields (4.3). \square

Remark 5.1. Note that the bound Ω given by (5.4) consists of two distinct terms. The first term αQ can be thought of as arising from an exponential time weighting of the cost, or, equivalently, from a uniform right shift of the open-loop dynamics ([1]). The second term $\alpha^{-1} \sum_{k=1}^p \alpha_k^2 \tilde{A}_k Q \tilde{A}_k^T$ arises naturally from a multiplicative white noise model ([3,4]). Such interpretations have no bearing on the results obtained here since only the bound Ω defined by (5.4) is required. Note that the bound is valid for all positive α .

Remark 5.2. The conservatism of the bound (5.4) is difficult to predict for two reasons. First, the overbounding (4.3) holds with respect to the partial ordering of the nonnegative-definite matrices for which no scalar measure of conservatism is available. And, second, the bound (4.3) is required to hold for all nonnegative-definite matrices Q and feedback gains $(D_{c1}, \dots, D_{cr}, E_{c1}, \dots, E_{cs})$. The conservatism will thus depend upon the actual values of $Q, D_{c1}, \dots, D_{cr}, E_{c1}, \dots, E_{cs}$ determined by solving (4.4).

6. The Auxiliary Minimization Problem and Necessary Conditions for Optimality

Rather than minimizing the actual cost (3.6), we shall consider the upper bound (4.6). This leads to the following problem.

Auxiliary Minimization Problem. Determine $(Q, D_{c1}, \dots, D_{cr}, E_{c1}, \dots, E_{cs})$ which minimizes

$$J(Q, D_{c1}, \dots, D_{cr}, E_{c1}, \dots, E_{cs}) \triangleq \text{tr } Q\tilde{R} \quad (6.1)$$

subject to

$$Q \in \mathbb{N}^n, \quad (6.2)$$

$$0 = \tilde{A}_a Q + Q \tilde{A}_a^T + \sum_{k=1}^p \gamma_k \tilde{A}_k Q \tilde{A}_k^T + \tilde{V}. \quad (6.3)$$

The relationship between the Auxiliary Minimization Problem and the Robust Stability and Performance Problem is straightforward as shown by the following observation.

Proposition 6.1. Suppose $(Q, D_{c1}, \dots, D_{cr}, E_{c1}, \dots, E_{cs})$ satisfies (6.2)–(6.4). Then

$$(\tilde{A} + \Delta\tilde{A}, \tilde{V}^{\frac{1}{2}}) \text{ is stabilizable for all variations in } \mathcal{U} \quad (6.4)$$

if and only if $\tilde{A} + \Delta\tilde{A}$ is asymptotically stable for all variations in \mathcal{U} . In this case,

$$J(D_{c1}, \dots, D_{cr}, E_{c1}, \dots, E_{cs}) \leq J(Q, D_{c1}, \dots, D_{cr}, E_{c1}, \dots, E_{cs}). \quad (6.5)$$

Proof. With Ω given by (5.4), Proposition 5.1 implies that (4.3) is satisfied. Since the hypotheses of Theorem 4.1 are satisfied, robust stability with performance bound (4.6) is guaranteed. Note that with definition (6.1), (6.5) is merely a restatement of (4.6). \square

The derivation of the necessary conditions for the Auxiliary Minimization Problem is based upon the Fritz John form of the Lagrange multiplier theorem.* Rigorous application of this technique requires that $(Q, D_{c1}, \dots, D_{cr}, E_{c1}, \dots, E_{cs})$ be restricted to the open set

$$S \triangleq \{(Q, D_{c1}, \dots, D_{cr}, E_{c1}, \dots, E_{cs}) : Q \in \mathbb{P}^n \text{ and } A \text{ is asymptotically stable}\},$$

* The Kuhn-Tucker theorem requires a priori verification of a constraint qualification which is difficult to confirm in the present context. The Fritz John version is less restrictive and hence more suitable.

where

$$A \triangleq \tilde{A}_\alpha \oplus \tilde{A}_\alpha + \sum_{k=1}^p \gamma_k \tilde{A}_k \otimes \tilde{A}_k.$$

The requirement $(Q, D_{c1}, \dots, D_{cr}, E_{c1}, \dots, E_{cs}) \in S$ implies that Q and its nonnegative-definite dual P are unique solutions to the modified Lyapunov equations (6.3) and

$$0 = \tilde{A}_\alpha^T P + P \tilde{A}_\alpha + \sum_{k=1}^p \gamma_k \tilde{A}_k^T P \tilde{A}_k + \tilde{R}. \quad (6.6)$$

An additional technical requirement is that $(Q, D_{c1}, \dots, D_{cr}, E_{c1}, \dots, E_{cs})$ be confined to the set

$$S^+ \triangleq \{(Q, D_{c1}, \dots, D_{cr}, E_{c1}, \dots, E_{cs}) \in S : \hat{C}_i Q \hat{C}_i^T > 0, \quad i = 1, \dots, r, \\ \text{and } \hat{B}_j^T P \hat{B}_j > 0, \quad j = 1, \dots, s\}.$$

The positive definiteness conditions in the definition of S^+ hold when \hat{C}_i and \hat{B}_j have full row and column rank, respectively, and Q and P are positive definite. As can be seen from the proof of Theorem 6.1 these conditions imply the existence of the projections ν_i and $\hat{\nu}_j$ corresponding to the two distinct types of feedback loops. Note that S^+ is open.

Remark 6.1. As pointed out in Remark 3.1, the set S may be empty in which case, of course, our results do not apply. As will be seen, however, our approach does not require explicit verification that S be nonempty since robust stability is obtained as a consequence of robust performance.

Remark 6.2. As will be seen, the constraint $(Q, D_{c1}, \dots, D_{cr}, E_{c1}, \dots, E_{cs}) \in S$ need not be verified in practice and is not required for either robust stability or robust performance since Proposition 6.1 shows that only (6.2)–(6.4) are needed. Rather, the set S constitutes sufficient conditions under which the Lagrange multiplier technique is applicable to the Auxiliary Minimization Problem. Specifically, the condition $Q > 0$ replaces (6.2) by an open set constraint, while the asymptotic stability of A serves as a normality condition which further implies that the dual P of Q is nonnegative definite.

Necessary conditions for the Auxiliary Minimization Problem can now be obtained.

Theorem 6.1. If $(Q, D_{c1}, \dots, D_{cr}, E_{c1}, \dots, E_{cs}) \in S^+$ solves the Auxiliary Minimization Problem with \mathcal{U} given by (5.1), then there exist $Q, P \in \mathbb{N}^n$ such that $D_{c1}, \dots, D_{cr}, E_{c1}, \dots, E_{cs}$ are given by

$$D_{ci} = -R_{\alpha i}^{-1} P_{\alpha i} Q \hat{C}_i^T (\hat{C}_i Q \hat{C}_i^T)^{-1}, \quad i = 1, \dots, r, \quad (6.7)$$

$$E_{cj} = -(\hat{B}_j^T P \hat{B}_j)^{-1} \hat{B}_j^T P Q_{\alpha j} V_{\alpha j}^{-1}, \quad j = 1, \dots, s, \quad (6.8)$$

and such that Q, P satisfy

$$\begin{aligned} 0 = & (A_\alpha - \sum_{i=1}^r B_i R_{\alpha i}^{-1} P_{\alpha i} \nu_i) Q + Q (A_\alpha - \sum_{i=1}^r B_i R_{\alpha i}^{-1} P_{\alpha i} \nu_i)^T + V_0 \\ & + \sum_{k=1}^p \gamma_k (A_k - \sum_{i=1}^r B_{ik} R_{\alpha i}^{-1} P_{\alpha i} \nu_i) Q (A_k - \sum_{i=1}^r B_{ik} R_{\alpha i}^{-1} P_{\alpha i} \nu_i)^T \\ & - \sum_{j=1}^s Q_{\alpha j} V_{\alpha j}^{-1} Q_{\alpha j}^T + \sum_{j=1}^s \hat{\nu}_{j\perp} Q_{\alpha j} V_{\alpha j}^{-1} Q_{\alpha j}^T \hat{\nu}_{j\perp}^T, \end{aligned} \quad (6.9)$$

$$\begin{aligned} 0 = & (A_\alpha - \sum_{j=1}^s \hat{\nu}_j Q_{\alpha j} V_{\alpha j}^{-1} C_j)^T P + P (A_\alpha - \sum_{j=1}^s \hat{\nu}_j Q_{\alpha j} V_{\alpha j}^{-1} C_j) + R_0 \\ & + \sum_{k=1}^p \gamma_k (A_k - \sum_{j=1}^s \hat{\nu}_j Q_{\alpha j} V_{\alpha j}^{-1} C_{jk})^T P (A_k - \sum_{j=1}^s \hat{\nu}_j Q_{\alpha j} V_{\alpha j}^{-1} C_{jk}) \\ & - \sum_{i=1}^r P_{\alpha i}^T R_{\alpha i}^{-1} P_{\alpha i} + \sum_{i=1}^r \nu_{i\perp}^T P_{\alpha i}^T R_{\alpha i}^{-1} P_{\alpha i} \nu_{i\perp}, \end{aligned} \quad (6.10)$$

$$\nu_i \triangleq Q \hat{C}_i^T (\hat{C}_i Q \hat{C}_i^T)^{-1} \hat{C}_i, \quad \nu_{i\perp} \triangleq I_n - \nu_i, \quad i = 1, \dots, r, \quad (6.11)$$

$$\hat{\nu}_j \triangleq \hat{B}_j (\hat{B}_j^T P \hat{B}_j)^{-1} \hat{B}_j^T P, \quad \hat{\nu}_{j\perp} \triangleq I_n - \hat{\nu}_j, \quad j = 1, \dots, s. \quad (6.12)$$

Furthermore, the auxiliary cost is given by

$$\begin{aligned} J(Q, D_{c1}, \dots, D_{cr}, E_{c1}, \dots, E_{cs}) = \\ \text{tr} \left[Q (R_0 + \sum_{i=1}^p \nu_i^T P_{\alpha i}^T R_{\alpha i}^{-1} R_i R_{\alpha i}^{-1} P_{\alpha i} \nu_i - 2 R_{0i} R_{\alpha i}^{-1} P_{\alpha i} \nu_i) \right]. \end{aligned} \quad (6.13)$$

Conversely, if there exist $Q, P \in \mathbb{R}^{n \times n}$ satisfying (6.9) and (6.10) then Q satisfies (6.3) with $(D_{c1}, \dots, D_{cr}, E_{c1}, \dots, E_{cs})$ given by (6.7) and (6.8), and $J(Q, D_{c1}, \dots, D_{cr}, E_{c1}, \dots, E_{cs})$ is given by (6.13).

Proof. To optimize (6.1) over the open set S^+ , subject to the constraint (6.3), form the Lagrangian

$$\mathcal{L}(Q, D_{c1}, \dots, D_{cr}, E_{c1}, \dots, E_{cs}) \triangleq \text{tr} [\lambda Q \tilde{R} + (\tilde{A} Q + Q \tilde{A}^T + \sum_{k=1}^p \gamma_k \tilde{A}_k Q \tilde{A}_k^T + \tilde{V}) P],$$

where the Lagrange multipliers $\lambda \geq 0$ and $P \in \mathbb{R}^{n \times n}$ are not both zero. Setting $\partial \mathcal{L} / \partial Q = 0$, $\lambda = 0$ implies $P = 0$ since A is asymptotically stable. Hence, without loss of generality set $\lambda = 1$. Thus the stationarity conditions are given by

$$\frac{\partial \mathcal{L}}{\partial Q} = \tilde{A}_\alpha^T P + P \tilde{A}_\alpha + \sum_{k=1}^p \gamma_k \tilde{A}_k^T P \tilde{A}_k + \tilde{R} = 0, \quad (6.14)$$

$$\frac{\partial \mathcal{L}}{\partial D_{ci}} = R_{ai} D_{ci} \hat{C}_i Q \hat{C}_i^T + P_{ai} Q \hat{C}_i^T = 0, \quad i = 1, \dots, r, \quad (6.15)$$

$$\frac{\partial \mathcal{L}}{\partial E_{cj}} = \hat{B}_j^T P \hat{B}_j E_{cj} V_{aj} + \hat{B}_j^T P Q_{aj} = 0, \quad j = 1, \dots, s. \quad (6.16)$$

Since $(Q, D_{c1}, \dots, D_{cr}, E_{c1}, \dots, E_{cs}) \in S^+$, $\hat{C}_i Q \hat{C}_i^T$ and $\hat{B}_j^T P \hat{B}_j$ are invertible and hence (6.15) and (6.16) imply (6.7) and (6.8). Finally, (6.9) and (6.10) are equivalent to (6.3) and (6.6). \square

Remark 6.3. Several special cases can be recovered formally from Theorem 6.1. For example, when the control weighting is nonsingular and the measurement noise is zero, i.e., when \hat{u}_i and y_i are absent for $i = 1, \dots, r$, delete (6.8) and set $\hat{v}_j = 0$ in (6.9). In this case the last two terms in (6.9) can be deleted. Deleting also the uncertainty terms A_k , B_{ik} , C_{jk} yields the results of [10] with the added features of correlated plant/measurement noise (V_{0j}) and cross weighting (R_{0i}). Furthermore, assuming a centralized structure for the static controller, i.e., $r = 1$, yields the usual static output feedback result ([8,9]).

7. Sufficient Conditions for Robust Stability and Performance

We now combine Proposition 6.1 and Theorem 6.1 to obtain sufficient conditions for robust stability and performance.

Theorem 7.1. Suppose there exist $Q, P \in \mathbb{N}^n$ satisfying (6.9) and (6.10). Then with $(D_{c1}, \dots, D_{cr}, E_{c1}, \dots, E_{cs})$ given by (6.6) and (6.7), $(\tilde{A} + \Delta \tilde{A}, \tilde{V}^{\frac{1}{2}})$ is stabilizable for all variations in \mathcal{U} if and only if $\tilde{A} + \Delta \tilde{A}$ is asymptotically stable for all variations in \mathcal{U} . In this case the performance of the closed-loop system satisfies the bound

$$J(D_{c1}, \dots, D_{cr}, E_{c1}, \dots, E_{cs}) \leq \text{tr} \left[Q(R_0 + \sum_{i=1}^p \nu_i^T P_{gi}^T R_{ai}^{-1} R_i R_{ai}^{-1} P_{ai} \nu_i - 2R_{0i} R_{ai}^{-1} P_{ai} \nu_i) \right]. \quad (7.1)$$

Proof. The converse of Theorem 6.1 shows that Q satisfies (6.3) with $(D_{c1}, \dots, D_{cr}, E_{c1}, \dots, E_{cs})$ given by (6.7) and (6.8). Hence, with the stabilizability assumption (6.4), Proposition 6.1 implies robust stability and performance. \square

Remark 7.1. The application of Theorem 7.1 in practice requires 1) numerical solution of (6.9) and (6.10), and 2) verification of the stabilizability hypothesis. No other assumptions need be verified in applying this result.

8. Concluding Remarks

We have developed a theory of robust decentralized output feedback via static control. The development permits the treatment of noisy and nonnoisy measurements, weighted and unweighted controls, and structured real-valued parameter uncertainties in the plant matrices. The theory provides a robustification of results given in [8-10] for both centralized and decentralized optimal output feedback. The theory is constructive in nature rather than existential. Specifically, the main result, Theorem 7.1, involves a coupled pair of modified Riccati equations (6.9), (6.10) whose solutions, when they exist, are used to explicitly construct feedback gains (6.7), (6.8) which are guaranteed to provide both robust stability and performance. Future research is required for evaluating the conservativeness of the theory. The numerical algorithms developed in [10] provide a starting point in this regard.

References

- [1] B.D.O. Anderson and J.B. Moore, *Linear Optimal Control* (Prentice-Hall, Englewood Cliffs, NJ, 1970).
- [2] D.S. Bernstein, The optimal projection equations for static and dynamic output feedback: The singular case, *IEEE Trans. Autom. Contr.*, Vol AC-32, pp. 1139-1143, 1987.
- [3] D.S. Bernstein, Robust static and dynamic output-feedback stabilization: Deterministic and stochastic perspectives, *IEEE Trans. Autom. Contr.*, Vol. AC-32, pp. 1076-1084, 1987.
- [4] D.S. Bernstein and D.C. Hyland, The optimal projection equations for reduced-order modelling, estimation and control of linear systems with multiplicative white noise, *J. Optim. Thy. Appl.*, Vol. 58, pp. 387-409, 1988.
- [5] J.W. Brewer, Kronecker products and matrix calculus in system theory, *IEEE Trans. Circ. Sys.*, Vol. CAS-25, pp. 772-781, 1978.
- [6] S.S.L. Chang and T.K.C. Peng, Adaptive guaranteed cost control of systems with uncertain parameters, *IEEE Trans. Autom. Contr.*, Vol. AC-17, pp. 474-483, 1972.
- [7] D.C. Hyland and D.S. Bernstein, The optimal projection equations for fixed-order dynamic compensation, *IEEE Trans. Autom. Contr.*, Vol. AC-29, pp. 1034-1037, 1984.
- [8] W.S. Levine and M. Athans, On the determination of the optimal constant output feedback gains for linear multivariable systems, *IEEE Trans. Autom. Contr.*, Vol. AC-15 pp. 44-48, 1970.
- [9] J. Medanic, On stabilization and optimization by output feedback, *Proc. Twelfth Asilomar Conf. Circ., Sys. Comp.*, pp. 412-416, 1978.
- [10] S. Renjen and D.P. Looze, Synthesis of decentralized output/state regulators, *Proc. Amer. Contr. Conf.*, Arlington, VA, pp. 758-762, 1982.
- [11] S.H. Wang and E.J. Davison, On the stabilization of decentralized control systems, *IEEE Trans. Autom. Contr.*, Vol. AC-18, pp. 473-478, 1973.
- [12] W.M. Wonham, *Linear Multivariable Control: A Geometric Approach* (Springer-Verlag, New York, 1979).

APPENDIX E: Singular Control

- 78. W. M. Haddad and D. S. Bernstein, "The Optimal Projection Equations for Reduced-Order State Estimation: The Singular Measurement Noise Case," IEEE Trans. Autom. Contr., Vol. AC-32, pp. 1135-1139, 1987.
- 79. D. S. Bernstein, "The Optimal Projection Equations for Static and Dynamic Output-Feedback: The Singular Case," IEEE Trans. Autom. Contr., Vol. AC-32, pp. 1139-1143, 1987.
- 130. Y. Halevi, W. M. Haddad, and D. S. Bernstein, "A Riccati Equation Approach to the Singular LQG Problem," (abstract), Amer. Contr. Conf., Pittsburgh, PA, June 1989.

The Optimal Projection Equations for Reduced-Order State Estimation: The Singular Measurement Noise Case

WASSIM M. HADDAD AND DENNIS S. BERNSTEIN

Dedicated to the memory of Professor Violet B. Haas
November 23, 1926–January 21, 1986

Abstract—The optimal projection equations for reduced-order state estimation are generalized to allow for singular (i.e., colored) measurement noise. The noisy and noise-free measurements serve as inputs to dynamic and static estimators, respectively. The optimal solution is characterized by necessary conditions which involve a pair of oblique projections corresponding to reduced estimator order and singular measurement noise intensity.

I. INTRODUCTION

It has recently been shown [1] that solutions to the steady-state reduced-order state-estimation problem can be characterized by means of a system of modified Riccati and Lyapunov equations coupled by an oblique projection. As in classical Kalman filter theory [2] however, this solution is based on the assumption that all measurements are corrupted by white noise. When the measurement noise is singular (i.e., colored), the optimal solution cannot be applied since the filter gains are given in terms of the

inverse of the noise intensity matrix. Hence, it is not surprising that a sizable body of literature has been devoted to the singular measurement noise problem in both continuous and discrete time [2]–[14]. For an overview of stochastic observer theory, see [15].

Much of the continuous-time singular estimation literature attempts to overcome the noise singularity by introducing new measurements obtained by differentiating noise-free measurements. The present note complements these results in the following way. For the available noisy and noise-free measurements we simultaneously design a reduced-order dynamic estimator for the noisy measurements and a static estimator for the noise-free measurements. We are not concerned here with the question of how the measurements are generated (e.g., via successive differentiation). Rather, our goal is to develop a unified dynamic/static estimation design theory which permits full utilization of both noisy and noise-free measurements. Application of these results to previously proposed approaches to singular estimation involving differentiation and transformation should be an interesting area for future research.

The results given herein directly generalize the results obtained in [1]. Specifically, the modified Riccati/Lyapunov equations are now coupled by a pair of oblique projections. As in [1] the requirement for reduced estimator order gives rise to the projection

$$\tau_2 = QP(QP)^* \quad (1.1)$$

where $()^*$ denotes group generalized inverse and Q and P are rank-deficient nonnegative-definite matrices analogous to the controllability and observability Gramians of the estimator. In addition, the presence of noise-free measurements

$$y(t) = C_2 x(t) \quad (1.2)$$

leads to the projection

$$\tau_1 = QC_2^T(C_2QC_2^T)^{-1}C_2 \quad (1.3)$$

where Q is the steady-state error covariance. The contribution of the present note is a concise, unified statement of the optimality conditions in a form which clearly displays the role of the oblique projections τ_1 and τ_2 in explicitly characterizing optimal static/dynamic (nonstrictly proper) estimators. An additional feature of the present note is the presence of state- and measurement-dependent white noise in the plant model. This model has been studied in a state-estimator context in [16]–[18] and has been justified as an approach to robustness in [19]–[22].

In Section III of the note, we consider the case in which the noisy and noise-free measurements are fed to the dynamic and static estimators, respectively. In Section IV, we note that feeding the noisy measurements to the static estimator results in an ill-posed problem, and we consider the general case in which the noise-free measurements are fed to both the static and dynamic estimators. Optimality conditions now lead to the interesting disjointness condition

$$0 = \tau_2 \tau_1 \quad (1.4)$$

concerning the relationship between the static and dynamic estimators. The meaning of (1.4) for proposed singular estimation schemes will be explored in future papers.

The goal of this note is confined to a rigorous development of necessary conditions for the optimal estimation problem. In support of this aim it should be noted that the usefulness of necessary conditions in optimization and optimal control has been amply demonstrated by classical results such as the maximum principle and Euler-Lagrange theory. For practical purposes, necessary conditions are largely free from restrictive special assumptions which invariably accompany sufficiency theory. Most importantly, success in addressing the problems of existence, sufficiency and global optimality is far more likely after the full elucidation of the necessary conditions has been achieved. Indeed, sufficiency conditions are often obtained by strengthening necessary conditions by means of additional restrictive assumptions.

Even without a complete resolution of questions pertaining to existence

Manuscript received August 7, 1986; revised November 14, 1986 and May 6, 1987. This work was supported in part by the Air Force Office of Scientific Research under Contract F49620-86-0002.

W. M. Haddad is with the Department of Mechanical Engineering, Florida Institute of Technology, Melbourne, FL 32901.

D. S. Bernstein is with Harris Corporation, Government Aerospace Systems Division, Melbourne, FL 32902.

IEEE Log Number 8716160

and sufficiency, the necessary conditions fulfill several immediate needs. Specifically, the structure of these conditions provides insight into the properties of the solution arising from optimality considerations. This has been demonstrated for the closely related problem of reduced-order modeling for which local minima are characterized in terms of an eigensystem decomposition [23]. Potentially more useful than insight for practical applications are prospects for constructing novel computational algorithms which avoid traditional gradient search methods. Thus far, two distinct algorithms have been developed, namely, an iterative method which exploits the structure of the oblique projection [23] and a homotopy algorithm which eliminates the need for eigensystem calculations and provides the means for attaining global optimality [24]. For computational purposes it should also be noted that under an existence assumption the necessary conditions are guaranteed to possess a solution to the problem, while sufficient conditions may fail in this regard.

III. PROBLEM STATEMENT AND MAIN THEOREM

Reduced-Order State-Estimation Problem

Given the n th-order observed system

$$\dot{x}(t) = \left(A + \sum_{i=1}^p v_i(t) A_i \right) x(t) + w_0(t), \quad (3.1)$$

$$y_i(t) = \left(C_i + \sum_{i=1}^p v_i(t) C_{i1} \right) x(t) + w_i(t), \quad (3.2)$$

$$y_2(t) = C_2 x(t), \quad (3.3)$$

II. NOTATION AND DEFINITIONS

$\mathbb{R}, \mathbb{R}^{r \times s}, \mathbb{R}^r, \mathbb{C}$

$I_n, (\cdot)^T, (\cdot)^*$

\oplus, \otimes

$\text{tr } Z$

τ, τ_1

n, l_1, l_2, n_e, p, q

\tilde{n}

x, x_e

y_1, y_2, y_e

$A, A_i; C_1, C_{i1}$

C_2

A_e, B_e, C_e, D_e

$v_i(t)$

$w_0(t), w_i(t)$

V_0

V_1

V_{01}

R

L

\tilde{A}, \tilde{A}_i

K

$\tilde{\tilde{A}}$

$\tilde{w}(t)$

\mathcal{P}

\tilde{R}

asymptotically stable matrix
nonnegative-semisimple matrix
nonnegative-definite matrix
positive-definite matrix

real numbers, $r \times s$ real matrices, $\mathbb{R}^{r \times 1}$, expectation

$n \times n$ identity, transpose, group generalized inverse [25, p. 124]

Kronecker sum, Kronecker product [26]

trace of a square matrix Z

$I_n - \tau, \tau \in \mathbb{R}^{n \times n}$

positive integers, $1 \leq n_e \leq n$

$n + n_e$

n, n_e -dimensional vectors

l_1, l_2, q -dimensional vectors

$n \times n$ matrices; $l_1 \times n$ matrices, $i = 1, \dots, p$

$l_2 \times n$ matrix

$n_e \times n_e, n_e \times l_1, q \times n_e, q \times l_2$ matrices

unit variance white noise, $i = 1, \dots, p$

n -dimensional, l_1 -dimensional white noise processes

$n \times n$ nonnegative-definite intensity of $w_0(t)$

$l_1 \times l_1$ positive-definite intensity of $w_1(t)$

$n \times l_1$ cross intensity of $w_0(t), w_1(t)$

$q \times q$ positive-definite matrix

$q \times n$ matrix

$$\begin{bmatrix} A & 0 \\ B_e C_1 & A_e \end{bmatrix}, \begin{bmatrix} A_i & 0 \\ B_e C_{i1} & 0 \end{bmatrix}, i = 1, \dots, p$$

$n_e \times l_2$ matrix

$$\begin{bmatrix} A & 0 \\ B_e C_1 + K C_2 & A_e \end{bmatrix}$$

$$\begin{bmatrix} w_0(t) \\ B_e w_1(t) \end{bmatrix}$$

$$\begin{bmatrix} V_0 & V_{01} B_e^T \\ B_e V_{01}^T & B_e V_1 B_e^T \end{bmatrix}$$

$$\begin{bmatrix} L^T R L - L^T R D_e C_2 - C_2^T D_e^T R L + C_2^T D_e^T R D_e C_2 & -L^T R C_e + C_2^T D_e^T R C_e \\ -C_2^T R L + C_2^T R D_e C_2 & C_2^T R C_e \end{bmatrix}$$

matrix with eigenvalues in open left-half plane
semisimple (nondefective) matrix with nonnegative eigenvalues
symmetric matrix with nonnegative eigenvalues
symmetric matrix with positive eigenvalues.

For arbitrary $n \times n$ Q , \tilde{Q} define:

$$V_{11} \triangleq V_1 + \sum_{i=1}^p C_{i1}(Q + \tilde{Q})C_{i1}^T,$$

$$Q \triangleq Q C_1^T + V_{01} + \sum_{i=1}^p A_i(Q + \tilde{Q})C_{i1}^T,$$

$$A_0 \triangleq A - Q V_{11}^{-1} C_1.$$

where $t \in [0, \infty)$, design an n_e th-order state estimator

$$\dot{x}_e(t) = A_e x_e(t) + B_e y_1(t), \quad (3.4)$$

$$y_e(t) = C_e x_e(t) + D_e y_2(t) \quad (3.5)$$

which minimizes the state-estimation error criterion

$$J(A_e, B_e, C_e, D_e) \triangleq \lim_{T \rightarrow \infty} \mathbb{E} [L x(T) - y_e(T)]^T R [L x(T) - y_e(T)]. \quad (3.6)$$

To guarantee that J is finite, assume that A is asymptotically stable and consider the set of asymptotically stable reduced-order (i.e., fixed-order) estimators

$$A \triangleq \{(A_r, B_r, C_r, D_r) : A_r \text{ is asymptotically stable}\}.$$

Since the value of J is independent of the internal realization of the transfer function corresponding to (3.4) and (3.5), without loss of generality we further restrict our attention to the set of admissible estimators

$$A^+ \triangleq \{(A_r, B_r, C_r, D_r) \in A : (A_r, B_r) \text{ is controllable and } (A_r, C_r) \text{ is observable}\}.$$

An additional technical requirement is that (A_r, B_r, C_r, D_r) be confined to the set

$$\hat{A}^+ \triangleq \{(A_r, B_r, C_r, D_r) \in A^+ : C_2(Q_1 - Q_{12}Q_2^{-1}Q_{12}^T)C_2^T \text{ is positive definite}\},$$

where

$$Q \triangleq \begin{bmatrix} Q_{11} & Q_{12} \\ Q_{12}^T & Q_{22} \end{bmatrix} \in \mathbb{R}^{A \times A}$$

satisfies

$$0 = A\bar{Q} + \bar{Q}A^T + \sum_{i=1}^p A_i \bar{Q} A_i^T + \bar{P}$$

and Q_2 is invertible since (A_r, B_r) is controllable. The positive definiteness condition holds when C_2 has full row rank and \bar{Q} is positive definite. As can be seen from the proof of Theorem 3.1, this condition implies the existence of the projection τ_1 defined below.

The following factorization lemma is needed for the statement of the main result.

Lemma 3.1: Suppose $n \times n$ \bar{Q} , \bar{P} are nonnegative definite. Then $\bar{Q}\bar{P}$ is nonnegative semisimple. If, in addition, $\text{rank } \bar{Q}\bar{P} = n_r$, then there exist $n_r \times n$ G , Γ and $n_r \times n_r$ invertible M such that

$$\bar{Q}\bar{P} = G^T M \Gamma, \quad (3.7a)$$

$$\Gamma G^T = I_{n_r}. \quad (3.7b)$$

Furthermore, G , M , and Γ are unique modulo a change of basis in \mathbb{R}^{n_r} .

Proof: The result follows from [27, Theorem 6.2.5]. \square

Since $\bar{Q}\bar{P}$ is semisimple (diagonalizable) it has a group generalized inverse $(\bar{Q}\bar{P})^+ = G^T M^{-1} \Gamma$ and

$$\tau_2 \triangleq \bar{Q}\bar{P}(\bar{Q}\bar{P})^+ = G^T \Gamma \quad (3.8)$$

is an oblique projection.

Theorem 3.1: Suppose A is asymptotically stable and $(A_r, B_r, C_r, D_r) \in \hat{A}^+$ solves the reduced-order state-estimation problem. Then there exist $n \times n$ nonnegative-definite matrices Q , \bar{Q} , and \bar{P} such that A_r , B_r , C_r , and D_r are given by

$$A_r = \Gamma(A - Q_r V_{11}^{-1} C_1) G^T, \quad (3.9)$$

$$B_r = \Gamma Q_r V_{11}^{-1}, \quad (3.10)$$

$$C_r = L \tau_{1\perp} G^T, \quad (3.11)$$

$$D_r = L Q C_2^T (C_2 Q C_2^T)^{-1} \quad (3.12)$$

and such that Q , \bar{Q} , and \bar{P} satisfy

$$0 = A\bar{Q} + \bar{Q}A^T + \sum_{i=1}^p A_i (\bar{Q} + \bar{Q}) A_i^T + V_0 - Q_r V_{11}^{-1} Q_r^T + \tau_{2\perp} Q_r V_{11}^{-1} Q_r^T \tau_{2\perp}^T, \quad (3.13)$$

$$0 = A\bar{Q} + \bar{Q}A^T + Q_r V_{11}^{-1} Q_r^T - \tau_{2\perp} Q_r V_{11}^{-1} Q_r^T \tau_{2\perp}^T, \quad (3.14)$$

$$0 = A^T \bar{Q} + \bar{P} A + \tau_{1\perp}^T L^T R L \tau_{1\perp} - \tau_{2\perp}^T \tau_{1\perp}^T L^T R L \tau_{1\perp} \tau_{2\perp}, \quad (3.15)$$

$$\text{rank } \bar{Q} = \text{rank } \bar{P} = \text{rank } \bar{Q}\bar{P} = n_r, \quad (3.16)$$

where

$$\tau_1 \triangleq Q C_2^T (C_2 Q C_2^T)^{-1} C_2. \quad (3.17)$$

Remark 3.1: Several special cases can be recovered from Theorem 3.1. For example, when the observation noise is nonsingular, i.e., when y_2 is absent, delete (3.12) and set $\tau_1 = 0$ [22]. Deleting also the multiplicative noise terms yields the Main Theorem of [1].

Specializing Theorem 3.1 to the full-order case $n_r = n$ reveals that the Lyapunov equation for \bar{P} is superfluous. In this case $G = \Gamma^{-1}$ and thus $G = \Gamma = I_n$ without loss of generality.

Corollary 3.1: Assume $n_r = n$, A is asymptotically stable and $(A_r, B_r, C_r, D_r) \in \hat{A}^+$ solves the full-order state-estimation problem. Then there exist $n \times n$ nonnegative-definite matrices Q and \bar{Q} such that A_r , B_r , C_r , D_r are given by

$$A_r = A - Q_r V_{11}^{-1} C_1, \quad (3.18)$$

$$B_r = Q_r V_{11}^{-1}, \quad (3.19)$$

$$C_r = L \tau_{1\perp}, \quad (3.20)$$

$$D_r = L Q C_2^T (C_2 Q C_2^T)^{-1} \quad (3.21)$$

and such that Q and \bar{Q} satisfy

$$0 = A\bar{Q} + \bar{Q}A^T + \sum_{i=1}^p A_i (\bar{Q} + \bar{Q}) A_i^T + V_0 - Q_r V_{11}^{-1} Q_r^T, \quad (3.22)$$

$$0 = A\bar{Q} + \bar{Q}A^T + Q_r V_{11}^{-1} Q_r^T. \quad (3.23)$$

Remark 3.2: Note that by setting $A_i = 0$, $C_i = 0$, $i = 1, \dots, p$, it follows that (3.22) and (3.23) are decoupled and (3.23) is superfluous. To recover the standard Kalman filter which involves nonsingular noise, set $C_2 = 0$, delete (3.21) and define $\tau_1 = 0$.

IV. ADDITIONAL ESTIMATOR PATHS

We now consider the more general estimator

$$\dot{x}_e(t) = A_e x_e(t) + B_e y_1(t) + K y_2(t), \quad (4.1)$$

$$y_e(t) = C_e x_e(t) + D_e y_1(t) + \bar{K} y_2(t) \quad (4.2)$$

involving the additional gains K and \bar{K} .

Note that the additional path introduced in (4.2) implies that J is infinite and thus the problem is meaningless. Hence, set $\bar{K} = 0$, and consider the additional path introduced by (4.1), i.e., filtering the noise-free measurement.

Replacing (3.4) by (4.1) and optimizing with respect to K yields

$$0 = G \bar{P} Q C_2^T, \quad (4.3)$$

which implies

$$0 = \tau_2 \tau_{1\perp}. \quad (4.4)$$

Using (4.3), $\bar{Q} = \tau_2 \bar{Q}$ and $\bar{P} = \bar{P} \tau_2$ [see (5.17)], the filter gains (3.9)–(3.15) become

$$A_r = \Gamma(A - Q_r V_{11}^{-1} C_1) G^T - K C_2 G^T, \quad (4.5)$$

$$B_r = \Gamma Q_r V_{11}^{-1}, \quad (4.6)$$

$$C_e = L\tau_{11}G^T, \quad (4.7)$$

$$D_e = LQC_2^T(C_2QC_2^T)^{-1}, \quad (4.8)$$

$$0 = A\bar{Q} + \bar{Q}A^T + \sum_{i=1}^r A_i(\bar{Q} + \bar{Q})A_i^T + V_0 - Q_e V_{12}^{-1} Q_e^T + \tau_{21} Q_e V_{12}^{-1} Q_e^T \tau_{21}^T, \quad (4.9)$$

$$0 = A\bar{Q} + \bar{Q}A^T + Q_e V_{12}^{-1} Q_e^T - \tau_{21} Q_e V_{12}^{-1} Q_e^T \tau_{21}^T - QC_2^T K^T G - G^T K C_2 Q_e, \quad (4.10)$$

$$0 = \bar{A}^T \bar{P} + \bar{P} \bar{A} + \tau_{11}^T L^T R L \tau_{11} - \tau_{21}^T \tau_{11}^T L^T R L \tau_{11} \tau_{21}, \quad (4.11)$$

where

$$\bar{A} \triangleq A - G^T K C_2, \quad \bar{A}_0 \triangleq A_0 - G^T K C_2.$$

V. PROOF OF THE MAIN THEOREM

Using the notation of Section II the augmented system (3.1)–(3.4) can be written as

$$\dot{x}(t) = \left(\bar{A} + \sum_{i=1}^r v_i(t) \bar{A}_i \right) x(t) + \bar{w}(t) \quad (5.1)$$

where

$$x(t) \triangleq [x^T(t), x_e^T(t)]^T.$$

To analyze (5.1) define the second-moment matrix

$$\bar{Q}(t) = \mathbb{E}[x(t)x^T(t)]. \quad (5.2)$$

It follows from [28, Theorem 8.5.5, p. 142] that $\bar{Q}(t)$ satisfies

$$\dot{\bar{Q}}(t) = \bar{A}\bar{Q}(t) + \bar{Q}(t)\bar{A}^T + \sum_{i=1}^r \bar{A}_i \bar{Q}(t) \bar{A}_i^T + \bar{V}, \quad t \geq 0. \quad (5.3)$$

Lemma 5.1: $A_e \in \mathcal{A}$ if and only if

$$\bar{\mathcal{A}} \triangleq \bar{A} \oplus \bar{A} + \sum_{i=1}^r \bar{A}_i \otimes \bar{A}_i$$

is asymptotically stable.

Proof: The result follows from properties of the Kronecker product applied to partitioned matrices. See [22], [26] for details. \square

Hence, $\bar{\mathcal{A}}$ stable assures

$$\bar{Q} \triangleq \lim_{t \rightarrow \infty} \mathbb{E}[x(t)x^T(t)]$$

exists. Furthermore, \bar{Q} and its nonnegative-definite dual \bar{P} are unique solutions of the modified Lyapunov equations

$$0 = \bar{A}\bar{Q} + \bar{Q}\bar{A}^T + \sum_{i=1}^r \bar{A}_i \bar{Q} \bar{A}_i^T + \bar{V}, \quad (5.4)$$

$$0 = \bar{A}^T \bar{P} + \bar{P} \bar{A} + \sum_{i=1}^r \bar{A}_i^T \bar{P} \bar{A}_i + \bar{R}. \quad (5.5)$$

Partition $\bar{n} \times \bar{n}$ \bar{Q} , \bar{P} into $n \times n$, $n \times n_e$, and $n_e \times n_e$ subblocks as

$$\bar{Q} = \begin{bmatrix} Q_{11} & Q_{12} \\ Q_{12}^T & Q_{22} \end{bmatrix}, \quad \bar{P} = \begin{bmatrix} P_{11} & P_{12} \\ P_{12}^T & P_{22} \end{bmatrix},$$

and define the $n \times n$ nonnegative-definite matrices

$$Q \triangleq Q_{11} - Q_{12} Q_{22}^{-1} Q_{12}^T, \quad P \triangleq P_{11} - P_{12} P_{22}^{-1} P_{12}^T,$$

$$\bar{Q} \triangleq Q_{12} Q_{22}^{-1} Q_{12}^T, \quad \bar{P} \triangleq P_{12} P_{22}^{-1} P_{12}^T$$

and the $n_e \times n$, $n_e \times n_e$, $n_e \times n$ matrices

$$G \triangleq Q_{22}^{-1} Q_{12}^T, \quad M \triangleq Q_{22} P_{22}, \quad \Gamma \triangleq -P_{22}^{-1} P_{12}^T.$$

To optimize (3.6) subject to the constraint (5.4) over $\bar{\mathcal{A}}^+$, form the Lagrangian

$$L(A_e, B_e, C_e, D_e, \bar{Q}, \bar{P}, \lambda)$$

$$\triangleq \text{tr} \left[\lambda \bar{Q} \bar{R} + \left(\bar{A} \bar{Q} + \bar{Q} \bar{A}^T + \sum_{i=1}^r \bar{A}_i \bar{Q} \bar{A}_i^T + \bar{V} \right) \bar{P} \right]$$

where the Lagrange multipliers $\lambda \geq 0$ and $\bar{P} \in \mathbb{R}^{\bar{n} \times \bar{n}}$ are not both zero and \bar{Q} and \bar{P} are viewed as arbitrary $\bar{n} \times \bar{n}$ matrix variables. Setting $\partial L / \partial \bar{Q} = 0$, $\lambda = 0$ implies $\bar{P} = 0$ since $(A_e, B_e, C_e, D_e) \in \bar{\mathcal{A}}^+$. Hence, without loss of generality, set $\lambda = 1$. Thus, the stationarity conditions are given by¹

$$\frac{\partial L}{\partial \bar{P}} = \bar{A} \bar{Q} + \bar{Q} \bar{A}^T + \sum_{i=1}^r \bar{A}_i \bar{Q} \bar{A}_i^T + \bar{V} = 0, \quad (5.6)$$

$$\frac{\partial L}{\partial \bar{Q}} = \bar{A}^T \bar{P} + \bar{P} \bar{A} + \sum_{i=1}^r \bar{A}_i^T \bar{P} \bar{A}_i + \bar{R} = 0, \quad (5.7)$$

$$\frac{\partial L}{\partial A_e} = P_{12}^T Q_{12} + P_{22} Q_{22} = 0, \quad (5.8)$$

$$\begin{aligned} \frac{\partial L}{\partial B_e} &= (P_{12}^T Q_{12} + P_{22} Q_{22}^T) C_1^T + \sum_{i=1}^r (P_{12}^T A_i C_1^T C_{1i}^T \\ &\quad + P_{22} B_i C_{1i} C_{1i}^T) + P_{12}^T V_{01} + P_{22} B_e V_{11} = 0, \end{aligned} \quad (5.9)$$

$$\frac{\partial L}{\partial C_e} = -R L Q_{12} + R D_e C_2 Q_{12} + R C_e Q_{22} = 0, \quad (5.10)$$

$$\frac{\partial L}{\partial D_e} = -R L Q_{12} C_2^T + R D_e C_2 Q_{12} C_2^T + R C_e Q_{12} C_2^T = 0. \quad (5.11)$$

Expanding (5.6) and (5.7) yields

$$0 = A Q_{11} + Q_{11} A^T + \sum_{i=1}^r A_i Q_{11} A_i^T + V_0, \quad (5.12)$$

$$0 = A Q_{12} + Q_{12} C_1^T B_e^T + Q_{12} A_e^T + \sum_{i=1}^r A_i Q_{12} C_{1i}^T B_e^T + V_{01} B_e^T, \quad (5.13)$$

$$0 = B_e C_1 Q_{12}^T + A_e Q_{22} + Q_{12}^T C_1^T B_e^T + Q_{22} A_e^T + \sum_{i=1}^r B_e C_{1i} Q_{12} C_{1i}^T B_e^T + B_e V_{11} B_e^T, \quad (5.14)$$

$$0 = A^T P_{12} + C_1^T B^T P_{22} + P_{12} A_e - L^T R C_e + C_1^T D_e^T R C_e, \quad (5.15)$$

$$0 = A_e^T P_{22} + P_{22} A_e + C_1^T R C_e. \quad (5.16)$$

Note that the (1, 1) subblock of (5.7) characterizing P_{11} has been omitted from the above equations since the estimator gains are independent of P_{11} .

Note that (5.8) implies (3.7a) and (3.7b). Since

$$Q_{22} P_{22} = P_{22}^{-1/2} (P_{22}^{1/2} Q_{22} P_{22}^{1/2}) P_{22}^{1/2},$$

M is positive semisimple. Sylvester's inequality yields (3.16). Note also that

$$\bar{Q} = \tau_2 \bar{Q}, \quad \bar{P} = \bar{P} \tau_2. \quad (5.17)$$

Next (3.10), (3.11), and (3.12) follow from (5.9), (5.10), and (3.11) by

¹ As shown in [29], the formula for the derivative of a scalar function with respect to symmetric arguments \bar{Q} and \bar{P} entails a modification of (5.6) and (5.7). Since these gradients are being set to zero, however, the final result is identical. Alternatively, \bar{Q} and \bar{P} can be viewed (as we are doing here) as arbitrary matrix variables. Symmetry is imposed only *a posteriori* by the form of (5.4) and (5.5) and the stability of \bar{A} . Hence, mathematically, the result of [29] is not required.

using the identities

$$Q_1 = Q + \hat{Q}, \quad P_1 = P + \hat{P}, \quad (5.18)$$

$$Q_{12} = \hat{Q}\Gamma^T, \quad P_{12} = -\hat{A}G^T, \quad (5.19)$$

$$Q_2 = \Gamma\hat{Q}\Gamma^T, \quad P_2 = G\hat{A}G^T. \quad (5.20)$$

Substituting (3.10), (3.11), (3.12) and (5.18)–(5.20) into (5.12)–(5.16) and using (5.12) + $G^T\Gamma(5.13)G - (5.13)G - (5.13G)^T$ and $G^T\Gamma(5.13)G - (5.13)G - (5.13G)^T$ yields (3.13) and (3.14). Using $\Gamma^TG(5.15)\Gamma - (5.15)\Gamma - (5.15\Gamma)^T$ yields (3.15). Finally, $\Gamma(5.13) - (5.14)$ or $G(5.15) - (5.16)$ yields (3.9). \square

Remark 5.1. Equations (4.5)–(4.11) are derived in a similar manner with \hat{A} replaced by \tilde{A} in (5.1).

REFERENCES

- [1] D. S. Bernstein and C. D. Hyland, "The optimal projection equations for reduced-order state estimation," *IEEE Trans. Automat. Contr.*, vol. AC-30, pp. 583–585, 1985.
- [2] H. Kwakernaak and R. Sivan, *Linear Optimal Control Systems*. New York: Wiley, 1972.
- [3] A. E. Bryson and D. E. Johansen, "Linear filtering for time varying systems using measurements containing colored noise," *IEEE Trans. Automat. Contr.*, vol. AC-10, pp. 4–10, 1965.
- [4] E. Fogel and Y. F. Huang, "Reduced order optimal state estimator for linear systems with partially noise-corrupted measurements," *IEEE Trans. Automat. Contr.*, vol. AC-25, pp. 994–996, 1980.
- [5] Y. T. Ju and V. B. Haas, "A duality principle for state estimation with partially noise-corrupted measurements," *Int. J. Contr.*, vol. 37, pp. 1039–1056, 1983.
- [6] U. Shaked, "Optimal nearly singular estimation of continuous linear stationary uniform rank systems," *Int. J. Contr.*, vol. 38, pp. 275–317, 1983.
- [7] V. B. Haas, "Reduced order state estimation for a linear system with exact measurements," *Automatica*, vol. 20, pp. 225–229, 1984.
- [8] U. Shaked, "Explicit solution to the singular discrete-time stationary linear filtering problem," *IEEE Trans. Automat. Contr.*, vol. AC-30, pp. 34–47, 1985.
- [9] V. B. Haas, "Minimal order Wiener filter for a system with exact measurements," *IEEE Trans. Automat. Contr.*, vol. AC-30, pp. 773–776, 1985.
- [10] J. M. Schumacher, "A geometric approach to the singular filtering problem," *IEEE Trans. Automat. Contr.*, vol. AC-30, pp. 1075–1082, 1985.
- [11] F. W. Fairman and L. Luk, "On reducing the order of Kalman filters for discrete-time stochastic systems having singular measurement noise," *IEEE Trans. Automat. Contr.*, vol. AC-30, pp. 1150–1152, 1985.
- [12] Y. Halevi and Z. J. Palmor, "Extended limiting forms of optimum observers and LQG regulators," *Int. J. Contr.*, vol. 43, pp. 193–212, 1986.
- [13] U. Shaked and B. Priel, "Explicit solutions to the singular discrete finite-time linear estimation problem," *Int. J. Contr.*, vol. 43, pp. 285–303, 1986.
- [14] W. M. Haddad and D. S. Bernstein, "The optimal projection equations for reduced-order, discrete-time state estimation for linear systems with multiplicative white noise," *Syst. Contr. Lett.*, vol. 8, pp. 381–388, 1987.
- [15] J. O'Reilly, *Observers for Linear Systems*. New York: Academic, 1983.
- [16] P. J. McLane, "Optimal linear filtering for linear systems with state-dependent noise," *Int. J. Contr.*, vol. 10, pp. 42–51, 1969.
- [17] R. B. Asher and C. S. Sims, "Reduced-order filtering with state dependent noise," in *Proc. Joint Amer. Contr. Conf.*, 1978.
- [18] M. J. Grimble, "Wiener and Kalman filters for systems with random parameters," *IEEE Trans. Automat. Contr.*, vol. AC-29, pp. 552–554, 1984.
- [19] D. S. Bernstein and D. C. Hyland, "The optimal projection/maximum entropy approach to designing low-order, robust controllers for flexible structures," in *Proc. 24th IEEE Conf. Dec. Contr.*, Fort Lauderdale, FL, Dec. 1985, pp. 745–752.
- [20] D. S. Bernstein and S. W. Greeley, "Robust controller synthesis using the maximum entropy design equations," *IEEE Trans. Automat. Contr.*, vol. AC-31, pp. 362–364, 1986.
- [21] D. S. Bernstein and S. W. Greeley, "Robust output-feedback stabilization: Deterministic and stochastic perspectives," in *Proc. American Contr. Conf.*, Seattle, WA, June 1986, pp. 1818–1826.
- [22] D. S. Bernstein and D. C. Hyland, "The optimal projection equations for reduced-order modelling, estimation and control of linear systems with multiplicative white noise," *J. Optimiz. Theory Appl.*, to be published.
- [23] D. C. Hyland and D. S. Bernstein, "The optimal projection equations for model reduction and the relationships among the methods of Wilson, Skelton and Moore," *IEEE Trans. Automat. Contr.*, vol. AC-30, pp. 1201–1211, 1985.
- [24] S. Richter, "A homotopy algorithm for solving the optimal projection equations for fixed-order dynamic compensation: Existence, convergence and global optimality," in *Proc. Amer. Contr. Conf.*, Minneapolis, MN, June 1987, pp. 1527–1531.
- [25] S. L. Campbell and C. D. Meyer, Jr., *Generalized Inverses of Linear Transformations*. London: Pitman, 1979.
- [26] J. W. Brewer, "Kronecker products and matrix calculus in system theory," *IEEE Trans. Circuits Syst.*, vol. CAS-25, pp. 772–781, 1978.
- [27] C. R. Rao and S. K. Mitra, *Generalized Inverse of Matrices and Its Applications*. New York: Wiley, 1971.
- [28] L. Arnold, *Stochastic Differential Equations: Theory and Applications*. New York: Wiley, 1974.
- [29] J. W. Brewer, "The gradient with respect to a symmetric matrix," *IEEE Trans. Automat. Contr.*, vol. AC-22, pp. 265–267, 1977.

The Optimal Projection Equations for Static and Dynamic Output Feedback: The Singular Case

DENNIS S. BERNSTEIN

Dedicated to the memory of Professor Violet B. Haas
November 23, 1926–January 21, 1986

Abstract—Oblique projections have been shown to arise naturally in both static and dynamic optimal design problems. For static controllers an oblique projection was inherent in the early work of Levine and Athans, while for dynamic controllers an oblique projection was developed by Hyland and Bernstein. This note is motivated by the following natural question: What is the relationship between the oblique projection arising in optimal static output feedback and the oblique projection arising in optimal fixed-order dynamic compensation? We show that in nonstrictly proper optimal output feedback there are, indeed, *three distinct oblique projections* corresponding to singular measurement noise, singular control weighting, and reduced compensator order. Moreover, we unify the Levine–Athans and Hyland–Bernstein approaches by rederiving the optimal projection equations for combined static/dynamic (nonstrictly proper) output feedback in a form which clearly illustrates the role of the three projections in characterizing the optimal feedback gains. Even when the dynamic component of the nonstrictly proper controller is of full order, the controller is characterized by four matrix equations which generalize the standard LQG result.

I. INTRODUCTION

The optimal static output-feedback problem [1], [2] and the optimal fixed-order dynamic-compensation problem [3], [4] have been extensively investigated. A salient feature of the necessary conditions for each of these problems is the presence of an oblique projection (idempotent matrix) which arises as a direct consequence of optimality. For the static problem with noise-free measurements (i.e., singular measurement noise) the necessary conditions involve the projection [2]

$$\tau_1 = QC^T(CQC^T)^{-1}C$$

where Q is the steady-state closed-loop state covariance. The dual projection

$$\tau_2 = B(B^T P B)^{-1} B^T P$$

arises analogously in the corresponding problem involving singular control weighting. Furthermore, for fixed-order dynamic compensation with noisy measurements, it has recently been shown [4] that the necessary conditions give rise to the projection

$$\tau_3 = \hat{Q}\hat{P}(\hat{Q}\hat{P})^{\#}$$

where $(\cdot)^{\#}$ denotes group generalized inverse and \hat{Q} and \hat{P} are rank-deficient nonnegative-definite matrices analogous to the controllability

Manuscript received February 6, 1986; revised November 25, 1986 and May 5, 1987. This work was supported in part by the Air Force Office of Scientific Research under Contracts AFOSR F49620-84-C-0015 and F49620-86-C-0002.

The author is with the Harris Corporation, Government Aerospace Systems Division, Melbourne, FL 32902.

IEEE Log Number 8716159.

and observability Gramians of the compensator. To understand the relationships among τ_1 , τ_2 , and τ_3 , the contribution of the present note is a unified treatment of the necessary conditions for optimal static/dynamic feedback compensation which clearly illustrates the role of the three projections in characterizing the optimal feedback gains. Even in the full-order case in which τ_3 is the identity, the result provides a generalization of the standard LQG result to nonstrictly proper controllers in which case the separation principle does not hold.

To clarify the ramifications of noise and weighting singularities in optimal output feedback, consider the problem of minimizing

$$J = \lim_{T \rightarrow \infty} \mathbb{E}[x^T R_0 x + u^T R_1 u] \quad (1.1)$$

with plant dynamics

$$\dot{x} = Ax + Bu + w_0, \quad (1.2)$$

$$y = Cx + w_1, \quad (1.3)$$

and nonstrictly proper feedback compensator

$$\dot{x}_c = A_c x_c + B_c y, \quad (1.4)$$

$$u = C_c x_c + D_c y. \quad (1.5)$$

As pointed out in [3], J is finite only if

$$0 = \text{tr}[D_c^T R_1 D_c V_1] \Leftrightarrow 0 = R_1 D_c V_1 \quad (1.6)$$

where V_1 denotes the intensity of w_1 . Clearly, when R_1 and V_1 are nonsingular (1.6) implies $D_c = 0$, and hence direct feedthrough is not permitted, i.e., the compensator must be strictly proper. Conversely, to utilize a static gain D_c , either R_1 or V_1 must be singular. By writing singular R_1 and V_1 without loss of generality as

$$R_1 = \begin{bmatrix} \tilde{R}_1 & 0 \\ 0 & 0 \end{bmatrix}, \quad V_1 = \begin{bmatrix} \tilde{V}_1 & 0 \\ 0 & 0 \end{bmatrix} \quad (1.7)$$

it follows that the static transmission between noisy measurements and weighted controls must be zero (see Fig. 1).

The reader will observe that three feedback paths which are not ruled

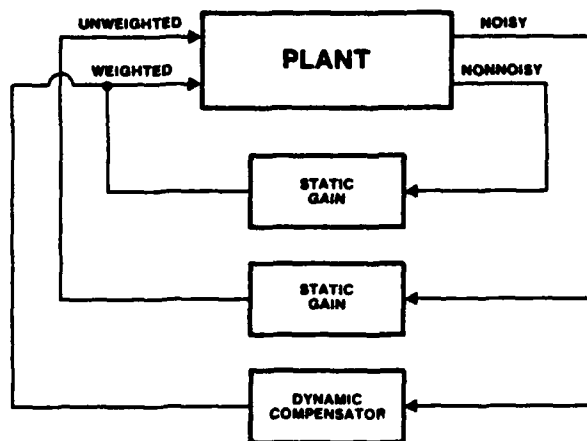


Fig. 1.

out by (1.6) do not appear in Fig. 1. Specifically: 1) nonnoisy measurements can be fed back to unweighted controls; 2) dynamic-compensator outputs can be fed back to unweighted controls; and 3) nonnoisy measurements can serve as inputs to the dynamic compensator. The reason for considering the more limited configuration shown in Fig. 1 is that only these paths are explicitly characterized by the necessary conditions. Hence, for simplicity we first consider only the scheme of Fig. 1, and later introduce the remaining permissible paths. Interestingly, while these additional gains are not completely determined by the necessary conditions, they appear to play an important role in governing geometric interrelationships among the three projections.

Two final comments are in order. First, since our results are carried out in a multiplicative noise setting, we generalize previous results on state feedback [15]–[18] and dynamic compensation [9]–[11]. The motivation for using a multiplicative white noise model is to represent plant parameter uncertainties and thereby obtain robust controllers [12]. Also, the derivations of the necessary conditions are straightforward extensions of the Lagrange multiplier technique used in [4] and hence have been omitted.

II. NOTATION AND DEFINITIONS

$\mathbb{R}, \mathbb{R}^{n \times s}, \mathbb{R}^r, \mathbb{E}$

$I_r, ()^T, ()^*$

\oplus, \otimes

τ_\perp

asymptotically stable matrix

nonnegative-semisimple matrix

$n, m_1, m_2, l_1, l_2, n_c, p$

$x, u_1, u_2, y_1, y_2, x_c$

$A, A_i; B_1, B_{1i}; C_1, C_{1i}$

B_2, C_2

A_c, B_c, C_c, D_c, E_c

$v_i(t)$

$w_0(t), w_1(t)$

V_0, V_1

V_{01}

R_0, R_1

R_{01}

\tilde{A}, \tilde{A}_i

\mathcal{P}, \mathcal{R}

\tilde{A}, \tilde{A}_i

real numbers, $r \times s$ real matrices, $\mathbb{R}^{r \times 1}$, expectation

$r \times r$ identity, transpose, group generalized inverse [13, p. 124]

Kronecker sum, Kronecker product

$I_n - \tau, \tau \in \mathbb{R}^{n \times n}$

matrix with eigenvalues in open left-half plane

semisimple (nondefective) matrix with nonnegative eigenvalues

positive integers

$n, m_1, m_2, l_1, l_2, n_c$ -dimensional vectors

$n \times n$ matrices, $n \times m_1$ matrices, $l_1 \times n$ matrices, $i = 1, \dots, p$

$n \times m_2$ matrix, $l_2 \times n$ matrix

$n_c \times n_c, n_c \times l_1, m_1 \times n_c, m_1 \times l_2, m_2 \times l_1$ matrices

unit variance white noise, $i = 1, \dots, p$

n -dimensional, l_1 -dimensional white noise

intensities of w_0, w_1 ; $V_0 \geq 0, V_1 > 0$

$n \times l_1$ cross intensity of w_0, w_1

state and control weightings; $R_0 \geq 0, R_1 > 0$

$n \times m_1$ cross weighting: $R_0 - R_{01} R_1^{-1} R_{01}^T \geq 0$

$A + B_1 D_c C_2 + B_2 E_c C_1, A_i + B_{1i} D_c C_2 + B_{2i} E_c C_{1i}, i = 1, \dots, p$

$$\begin{bmatrix} I_n \\ (B_2 E_c)^T \end{bmatrix}^T \begin{bmatrix} V_0 & V_{01} \\ V_{01}^T & V_1 \end{bmatrix} \begin{bmatrix} I_n \\ (B_2 E_c)^T \end{bmatrix} \cdot \begin{bmatrix} I_n \\ D_c C_2 \end{bmatrix}^T \begin{bmatrix} R_0 & R_{01} \\ R_{01}^T & R_1 \end{bmatrix} \begin{bmatrix} I_n \\ D_c C_2 \end{bmatrix} \\ \begin{bmatrix} \tilde{A} & B_1 C_c \\ B_c C_1 & A_c \end{bmatrix} \cdot \begin{bmatrix} \tilde{A}_i & B_{1i} C_c \\ B_{ci} C_{1i} & 0 \end{bmatrix}$$

\bar{V}
 \bar{R}

$$\begin{bmatrix} \bar{V} & V_0 B_c^T + B_2 E_c V_1 B_c^T \\ B_c V_0^T + B_2 V_1 (B_2 E_c)^T & B_c V_1 B_c^T \end{bmatrix}$$

$$\begin{bmatrix} \bar{R} & R_0 C_c + (D_c C_2)^T R_1 C_c \\ C_c^T R_0^T + C_c^T R_1 D_c C_2 & C_c^T R_1 C_c \end{bmatrix}$$

For arbitrary $n \times n$ $Q, P, \bar{Q}, \bar{P}, \tau_1, \tau_2$ define:

$$R_{1i} \triangleq R_1 + \sum_{j=1}^p B_{1j}^T P B_{1j}, \quad V_{1i} \triangleq V_1 + \sum_{j=1}^p C_{1j} Q C_{1j}^T,$$

$$Q_{1i} \triangleq Q C_1^T + V_0 + \sum_{j=1}^p A_{1j} Q C_{1j}^T, \quad \bar{Q}_{1i} \triangleq \bar{Q} \bar{C}_1^T + \bar{V}_0 + \sum_{j=1}^p \bar{A}_{1j} \bar{Q} \bar{C}_{1j}^T,$$

$$\bar{R}_{1i} \triangleq \bar{R}_1 + \sum_{j=1}^p \bar{B}_{1j}^T (\bar{P} + \bar{P}) \bar{B}_{1j}, \quad \bar{V}_{1i} \triangleq \bar{V}_1 + \sum_{j=1}^p \bar{C}_{1j} (\bar{Q} + \bar{Q}) \bar{C}_{1j}^T,$$

$$\bar{Q}_{1i} \triangleq \bar{Q} \bar{C}_1^T + \bar{V}_0 + \sum_{j=1}^p \bar{A}_{1j} (\bar{Q} + \bar{Q}) \bar{C}_{1j}^T,$$

$$\bar{P}_{1i} \triangleq \bar{P}_1 P + \bar{R}_{1i} + \sum_{j=1}^p \bar{B}_{1j}^T (\bar{P} + \bar{P}) \bar{B}_{1j},$$

$$\bar{A}_{1i} \triangleq \bar{A} - \bar{B}_1 \bar{R}_{1i}^{-1} (\bar{Q}_{1i} + \bar{B}_1^T \bar{P}) \tau_1 - \tau_2 (\bar{Q}_{1i} + \bar{Q} \bar{C}_1^T) \bar{V}_{1i}^{-1} \bar{C}_{1i},$$

$$\bar{A}_{1i} \triangleq \bar{A}_i - \bar{B}_{1i} \bar{R}_{1i}^{-1} (\bar{Q}_{1i} + \bar{B}_{1i}^T \bar{P}) \tau_1 - \tau_2 (\bar{Q}_{1i} + \bar{Q} \bar{C}_1^T) \bar{V}_{1i}^{-1} \bar{C}_{1i},$$

$$A_{Q_i} \triangleq \bar{A} - (\tau_{1i} \bar{Q}_{1i} - \tau_2 \bar{Q} \bar{C}_1^T) \bar{V}_{1i}^{-1} \bar{C}_{1i}, \quad \bar{A}_{1i} \triangleq \bar{A} - \bar{B}_1 \bar{R}_{1i}^{-1} (\bar{Q}_{1i} + \bar{B}_1^T \bar{P}) \tau_1,$$

$$\bar{R}_0 \triangleq R_0 - R_0 \bar{R}_{1i}^{-1} (\bar{Q}_{1i} + \bar{B}_1^T \bar{P}) \tau_1 - \tau_2^T (\bar{Q}_{1i} + \bar{B}_1^T \bar{P}) \bar{R}_{1i}^{-1} \bar{R}_0^T$$

$$+ \tau_1^T (\bar{Q}_{1i} + \bar{B}_1^T \bar{P}) \bar{R}_{1i}^{-1} \bar{R}_1 \bar{R}_{1i}^{-1} (\bar{Q}_{1i} + \bar{B}_1^T \bar{P}) \tau_1,$$

$$\bar{V}_0 \triangleq V_0 - V_0 \bar{V}_{1i}^{-1} (\bar{Q}_{1i} + \bar{Q} \bar{C}_1^T) \tau_2 - \tau_2 (\bar{Q}_{1i} + \bar{Q} \bar{C}_1^T) \bar{V}_{1i}^{-1} V_0^T$$

$$+ \tau_2 (\bar{Q}_{1i} + \bar{Q} \bar{C}_1^T) \bar{V}_{1i}^{-1} V_1 \bar{V}_{1i}^{-1} (\bar{Q}_{1i} + \bar{Q} \bar{C}_1^T) \tau_2^T.$$

III. STATIC OUTPUT FEEDBACK

Static Output Feedback Problem

Given the controlled system

$$\dot{x}(t) = \left(A + \sum_{i=1}^p u_i(t) A_i \right) x(t)$$

$$+ \left(B_1 + \sum_{i=1}^p u_i(t) B_{1i} \right) u_1(t) + B_2 u_2(t) + w_0(t), \quad (3.1)$$

$$y_1(t) = \left(C_1 + \sum_{i=1}^p u_i(t) C_{1i} \right) x(t) + w_1(t), \quad (3.2)$$

$$y_2(t) = C_2 x(t) \quad (3.3)$$

where $t \in [0, \infty)$, determine D_c and E_c such that the static output feedback law

$$u_1(t) = D_c y_2(t), \quad (3.4)$$

$$u_2(t) = E_c y_1(t) \quad (3.5)$$

minimizes the performance criterion

$$J \triangleq \lim_{t \rightarrow \infty} \mathbb{E} [x(t)^T R_0 x(t) + 2x(t)^T R_0 u_1(t) + u_1(t)^T R_1 u_1(t)]. \quad (3.6)$$

To develop necessary conditions for this problem, D_c and E_c must be restricted to the set of second-moment-stabilizing gains

$$\mathcal{S} \triangleq \left\{ (D_c, E_c) : \bar{A} \oplus \bar{A} + \sum_{i=1}^p \bar{A}_i \oplus \bar{A}_i \text{ is asymptotically stable} \right\}.$$

The requirement $(D_c, E_c) \in \mathcal{S}$ implies the existence of the steady-state closed-loop state covariance $Q \triangleq \lim_{t \rightarrow \infty} \mathbb{E} [x(t)x(t)^T]$. Furthermore, Q and its nonnegative-definite dual P are the unique solutions of the modified Lyapunov equations

$$0 = \bar{A}Q + Q\bar{A}^T + \sum_{i=1}^p \bar{A}_i Q \bar{A}_i^T + \bar{V}, \quad (3.7)$$

$$0 = \bar{A}^T P + P\bar{A} + \sum_{i=1}^p \bar{A}_i^T P \bar{A}_i + \bar{R}. \quad (3.8)$$

An additional technical assumption is that (D_c, E_c) be confined to the set

$$\mathcal{S}^+ \triangleq \{(D_c, E_c) \in \mathcal{S} : C_2 Q C_2^T > 0 \text{ and } B_2^T P B_2 > 0\}.$$

In order to obtain closed-form expressions for the feedback gains we make the additional assumption here and in Section IV that

$$[B_{1i} \neq 0 = C_{1i} = 0], \quad i = 1, \dots, p, \quad (3.9)$$

i.e., for each i , B_{1i} and C_{1i} are not both nonzero. By optimizing (3.6) with respect to D_c and E_c and manipulating (3.7) and (3.8), we obtain the following result.

Theorem 3.1: Suppose $(D_c, E_c) \in \mathcal{S}^+$ solves the static output feedback problem. Then there exist $n \times n$ nonnegative-definite Q, P such that

$$D_c = -R_{1i}^{-1} \bar{Q}_{1i} Q C_2^T (C_2 Q C_2^T)^{-1}, \quad (3.10)$$

$$E_c = -(B_2^T P B_2)^{-1} B_2^T P Q_{1i} V_{1i}^{-1}, \quad (3.11)$$

and such that Q and P satisfy

$$0 = (A - B_1 R_{1i}^{-1} \bar{Q}_{1i} \tau_1) Q + Q (A - B_1 R_{1i}^{-1} \bar{Q}_{1i} \tau_1)^T + V_0$$

$$+ \sum_{i=1}^p (A_i - B_{1i} R_{1i}^{-1} \bar{Q}_{1i} \tau_1) Q (A_i - B_{1i} R_{1i}^{-1} \bar{Q}_{1i} \tau_1)^T$$

$$- Q_{1i} V_{1i}^{-1} Q_{1i}^T + \tau_{2i} Q_{1i} V_{1i}^{-1} Q_{1i}^T \tau_{2i}^T, \quad (3.12)$$

$$0 = (A - \tau_2 Q_{1i} V_{1i}^{-1} C_{1i})^T P + P (A - \tau_2 Q_{1i} V_{1i}^{-1} C_{1i}) + R_0$$

$$+ \sum_{i=1}^p (A_i - \tau_2 Q_{1i} V_{1i}^{-1} C_{1i})^T P (A_i - \tau_2 Q_{1i} V_{1i}^{-1} C_{1i})$$

$$- \bar{Q}_{1i}^T R_{1i}^{-1} \bar{Q}_{1i} + \tau_{1i}^T \bar{Q}_{1i}^T R_{1i}^{-1} \bar{Q}_{1i} \tau_{1i}, \quad (3.13)$$

where

$$\tau_1 \triangleq Q C_2^T (C_2 Q C_2^T)^{-1} C_2, \quad (3.14)$$

$$\tau_2 \triangleq B_2 (B_2^T P B_2)^{-1} B_2^T P. \quad (3.15)$$

Remark 3.1: Several special cases can be recovered formally from Theorem 3.1. For example, when the control weighting is nonsingular

and the measurement noise is zero, i.e., when u_2 and y_1 are absent, delete (3.11) and set $\tau_2 = 0$. Deleting also the multiplicative noise terms yields the usual static output feedback result [1], [2].

IV. DYNAMIC OUTPUT FEEDBACK

We now expand the formulation of the static problem to include a purely dynamic (strictly proper) dynamic compensator.

Dynamic Output Feedback Problem

Given (3.1)–(3.3), determine A_c, B_c, C_c, D_c, E_c such that the static and dynamic output feedback law

$$\dot{x}_c(t) = A_c x_c(t) + B_c y_1(t), \quad (4.1)$$

$$u_1(t) = C_c x_c(t) + D_c y_2(t), \quad (4.2)$$

$$u_2(t) = E_c y_1(t) \quad (4.3)$$

minimizes the performance criterion (3.6).

We restrict our attention to second-moment-stabilizing controllers

$$\mathcal{D} \triangleq \left\{ (A_c, B_c, C_c, D_c, E_c); \tilde{A} \oplus \tilde{A} + \sum_{i=1}^p \tilde{A}_i \oplus \tilde{A}_i \text{ is asymptotically stable and } (A_c, B_c, C_c) \text{ is minimal} \right\},$$

which implies the existence of $\tilde{Q} \triangleq \lim_{t \rightarrow \infty} \mathbb{E}[\tilde{x}(t)\tilde{x}(t)^T]$, where $\tilde{x}(t) \triangleq (x(t)^T, x_c(t)^T)^T$. Furthermore, \tilde{Q} and its dual \tilde{P} are the unique solutions of the modified Lyapunov equations

$$0 = \tilde{A}\tilde{Q} + \tilde{Q}\tilde{A}^T + \sum_{i=1}^p \tilde{A}_i \tilde{Q} \tilde{A}_i^T + \tilde{V}, \quad (4.4)$$

$$0 = \tilde{A}^T \tilde{P} + \tilde{P} \tilde{A} + \sum_{i=1}^p \tilde{A}_i^T \tilde{P} \tilde{A}_i + \tilde{R}. \quad (4.5)$$

Partitioning

$$\tilde{Q} = \begin{bmatrix} Q_{11} & Q_{12} \\ Q_{12}^T & Q_{22} \end{bmatrix}, \quad \tilde{P} = \begin{bmatrix} P_{11} & P_{12} \\ P_{12}^T & P_{22} \end{bmatrix},$$

where Q_{12} and P_{12} are $n \times n_c$, we also require

$$\mathcal{D}^* \triangleq \{(A_c, B_c, C_c, D_c, E_c) \in \mathcal{D} : C_c(Q_{11} - Q_{12}Q_{22}^{-1}Q_{12}^T)C_c^T > 0 \text{ and } B_c^T(P_{11} - P_{12}P_{22}^{-1}P_{12}^T)B_c > 0\}.$$

Optimizing (3.6) over \mathcal{D}^* , introducing new variables

$$Q \triangleq Q_{11} - Q_{12}Q_{22}^{-1}Q_{12}^T, \quad P \triangleq P_{11} - P_{12}P_{22}^{-1}P_{12}^T, \quad (4.6)$$

$$\tilde{Q} \triangleq Q_{12}Q_{22}^{-1}Q_{12}^T, \quad \tilde{P} \triangleq P_{12}P_{22}^{-1}P_{12}^T, \quad (4.7)$$

and manipulating (4.4) and (4.5), we obtain the dynamic extension of Theorem 3.1. The following lemma is required for the statement of the result.

Lemma 4.1: Suppose $n \times n$ \tilde{Q}, \tilde{P} are nonnegative definite. Then $\tilde{Q}\tilde{P}$ is nonnegative semisimple. If, in addition, $\text{rank } \tilde{Q}\tilde{P} = n_c$, then there exist $n_c \times n$ G, Γ and $n_c \times n_c$ invertible M such that

$$\tilde{Q}\tilde{P} = G^T M \Gamma, \quad \Gamma G^T = I_{n_c}. \quad (4.8a, b)$$

Proof: The result follows from [14, Theorem 6.2.5].

Since $\tilde{Q}\tilde{P}$ is semisimple it has a group inverse $(\tilde{Q}\tilde{P})^\# = G^T M^{-1} \Gamma$ and

$$\tau_3 \triangleq \tilde{Q}\tilde{P}(\tilde{Q}\tilde{P})^\# = G^T \Gamma \quad (4.9)$$

is an oblique projection.

Theorem 4.1: Suppose $(A_c, B_c, C_c, D_c, E_c) \in \mathcal{D}^*$ solves the dynamic output feedback problem. Then there exist $n \times n$ nonnegative-definite $Q,$

P, \tilde{Q}, \tilde{P} such that

$$A_c = \Gamma(A - B_1 \tilde{R}_{11}^{-1} \tilde{\Phi}_s - \tilde{Q}_s \tilde{P}_{11}^{-1} C_1) G^T, \quad (4.10)$$

$$B_c = \Gamma(\tau_{21} \tilde{Q}_s - \tau_2 \tilde{Q} C_1^T) \tilde{P}_{11}^{-1}, \quad (4.11)$$

$$C_c = -\tilde{R}_{11}^{-1} (\tilde{\Phi}_s \tau_{11} - B_1^T \tilde{P}_{11}) G^T, \quad (4.12)$$

$$D_c = -\tilde{R}_{11}^{-1} (\tilde{\Phi}_s + B_1^T \tilde{P}) Q C_1^T (C_1 Q C_1^T)^{-1}, \quad (4.13)$$

$$E_c = -(B_2^T P B_2)^{-1} B_2^T P (\tilde{Q}_s + \tilde{Q} C_1^T) \tilde{P}_{11}^{-1} \quad (4.14)$$

where τ_1 and τ_2 are given by (3.14) and (3.15), G, Γ satisfy (4.8a), (4.8b), and such that, with τ_3 given by (4.9), $Q, P, \tilde{Q},$ and \tilde{P} satisfy

$$\begin{aligned} 0 = & \tilde{A} Q + Q \tilde{A}^T + \tilde{P}_0 + \sum_{i=1}^p [\tilde{A}_i Q \tilde{A}_i^T + (\tilde{A}_i - B_{1i} \tilde{R}_{1i}^{-1} [\tilde{\Phi}_s \tau_{11} - B_1^T \tilde{P}_{11}]) \\ & \cdot \tilde{Q} (\tilde{A}_i - B_{1i} \tilde{R}_{1i}^{-1} [\tilde{\Phi}_s \tau_{11} - B_1^T \tilde{P}_{11}])^T - (\tau_{21} \tilde{Q}_s - \tau_2 \tilde{Q} C_1^T) \tilde{P}_{11}^{-1} \\ & \cdot (\tau_{21} \tilde{Q}_s - \tau_2 \tilde{Q} C_1^T)^T + \tau_{31} (\tau_{21} \tilde{Q}_s - \tau_2 \tilde{Q} C_1^T) \tilde{P}_{11}^{-1} (\tau_{21} \tilde{Q}_s - \tau_2 \tilde{Q} C_1^T)^T \tau_{31}^T], \end{aligned} \quad (4.15)$$

$$\begin{aligned} 0 = & \tilde{A}^T P + P \tilde{A} + \tilde{R}_0 + \sum_{i=1}^p [\tilde{A}_i^T P \tilde{A}_i + (\tilde{A}_i - [\tau_{21} \tilde{Q}_s - \tau_2 \tilde{Q} C_1^T] \tilde{P}_{11}^{-1} C_{1i})^T \\ & \cdot \tilde{P} (\tilde{A}_i - [\tau_{21} \tilde{Q}_s - \tau_2 \tilde{Q} C_1^T] \tilde{P}_{11}^{-1} C_{1i}) - (\tilde{\Phi}_s \tau_{11} - B_1^T \tilde{P}_{11})^T \tilde{R}_{11}^{-1} \\ & \cdot (\tilde{\Phi}_s \tau_{11} - B_1^T \tilde{P}_{11}) + \tau_{31}^T (\tilde{\Phi}_s \tau_{11} - B_1^T \tilde{P}_{11})^T \tilde{R}_{11}^{-1} (\tilde{\Phi}_s \tau_{11} - B_1^T \tilde{P}_{11}) \tau_{31}], \end{aligned} \quad (4.16)$$

$$\begin{aligned} 0 = & A_m \tilde{Q} + \tilde{Q} A_m^T + (\tau_{21} \tilde{Q}_s - \tau_2 \tilde{Q} C_1^T) \tilde{P}_{11}^{-1} (\tau_{21} \tilde{Q}_s - \tau_2 \tilde{Q} C_1^T)^T \\ & - \tau_{31} (\tau_{21} \tilde{Q}_s - \tau_2 \tilde{Q} C_1^T) \tilde{P}_{11}^{-1} (\tau_{21} \tilde{Q}_s - \tau_2 \tilde{Q} C_1^T)^T \tau_{31}^T, \end{aligned} \quad (4.17)$$

$$\begin{aligned} 0 = & A_m^T \tilde{P} + \tilde{P} A_m + (\tilde{\Phi}_s \tau_{11} - B_1^T \tilde{P}_{11})^T \tilde{R}_{11}^{-1} (\tilde{\Phi}_s \tau_{11} - B_1^T \tilde{P}_{11}) \\ & - \tau_{31}^T (\tilde{\Phi}_s \tau_{11} - B_1^T \tilde{P}_{11})^T \tilde{R}_{11}^{-1} (\tilde{\Phi}_s \tau_{11} - B_1^T \tilde{P}_{11}) \tau_{31}, \end{aligned} \quad (4.18)$$

$$\text{rank } \tilde{Q} = \text{rank } \tilde{P} = \text{rank } \tilde{Q}\tilde{P} = n_c. \quad (4.19)$$

Remark 4.1: Setting $\tau_1 = \tau_2 = 0, D_c = 0, E_c = 0$ yields the results of [4], [11].

Remark 4.2: Suppose $n_c = n$ so that $\tau_3 = I_n$. Then the resulting full-order nonstrictly proper controller is characterized by four matrix equations which generalize the standard LQG result. In this case the separation principle is no longer valid.

V. ADDITIONAL FEEDBACK PATHS

We now introduce the feedback paths not shown in Fig. 1. For the static problem replace (3.5) by

$$u_2(t) = E_c y_1(t) + K_1 y_2(t). \quad (5.1)$$

Optimizing with respect to K_1 yields the additional condition

$$0 = C_1 Q P B_2 \quad (5.2)$$

which implies

$$0 = \tau_2 \tau_1. \quad (5.3)$$

This geometric condition holds when K_1 is optimally chosen. Although K_1

is not given explicitly, it does play a role in the necessary conditions since A is replaced by $A + B_2 K_1 C_2$.

For the dynamic problem replace (4.1) and (4.3) by

$$\dot{x}_c(t) = A_c x_c(t) + B_c y_1(t) + K_3 y_2(t), \quad (5.4)$$

$$u_2(t) = E_c y_1(t) + K_2 x_c(t) + K_1 y_2(t). \quad (5.5)$$

Optimizing with respect to K_1, K_2, K_3 yields

$$0 = C_2(QP + \bar{Q}P + Q\bar{P})B_2, \quad (5.6a)$$

$$0 = \bar{Q}PB_2, \quad (5.6b)$$

$$0 = C_2Q\bar{P}, \quad (5.6c)$$

which imply

$$0 = \tau_2 \tau_1, \quad (5.7a)$$

$$0 = \tau_2 \tau_3, \quad (5.7b)$$

$$0 = \tau_3 \tau_1. \quad (5.7c)$$

Note that (5.7b) and (5.7c) imply

$$0 = \tau_2 \tau_3 \tau_1. \quad (5.8)$$

Using (5.7b), (5.7c), $\bar{Q} = \tau_3 \bar{Q}, \bar{P} = \bar{P} \tau_3$ (see [4]), (4.10)–(4.18) become

$$A_c = \Gamma(A - B_1 \bar{R}_{12}^{-1} \bar{\Phi}_s - \bar{Q}_s \bar{V}_{12}^{-1} C_1 + B_2 K_1 C_2)G^T + \Gamma B_2 K_2 - K_3 C_2 G^T, \quad (5.9)$$

$$B_c = \Gamma \tau_2 \bar{Q}_s \bar{V}_{12}^{-1}, \quad (5.10)$$

$$C_c = -\bar{R}_{12}^{-1} \bar{\Phi}_s \tau_{12} G^T, \quad (5.11)$$

$$D_c = -\bar{R}_{12}^{-1} \bar{\Phi}_s Q C_2^T (C_2 Q C_2^T)^{-1}, \quad (5.12)$$

$$E_c = -(B_2^T P B_2)^{-1} B_2^T P \bar{Q}_s \bar{V}_{12}^{-1}, \quad (5.13)$$

$$\begin{aligned} 0 = & \bar{A}_1 Q + Q \bar{A}_1^T + \bar{V}_0 + \sum_{i=1}^p [\bar{A}_i Q \bar{A}_i^T + (\bar{A}_i - B_{1i} \bar{R}_{1i}^{-1} \bar{\Phi}_{s1i}) \\ & \cdot \bar{Q}(\bar{A}_i - B_{1i} \bar{R}_{1i}^{-1} \bar{\Phi}_{s1i})^T] - \tau_{21} \bar{Q}_s \bar{V}_{12}^{-1} \bar{Q}_s^T \tau_{21}^T \\ & + \tau_{31} \tau_{21} \bar{Q}_s \bar{V}_{12}^{-1} \bar{Q}_s^T \tau_{21}^T \tau_{31}^T, \end{aligned} \quad (5.14)$$

$$\begin{aligned} 0 = & \bar{A}_2^T P + P \bar{A}_2 + \bar{R}_0 + \sum_{i=1}^p [\bar{A}_i^T P \bar{A}_i + (\bar{A}_i + \tau_2 \bar{Q}_s \bar{V}_{12}^{-1} C_{1i})^T \\ & \cdot \bar{P}(\bar{A}_i + \tau_2 \bar{Q}_s \bar{V}_{12}^{-1} C_{1i})] - \tau_{11}^T \bar{\Phi}_s^T \bar{R}_{11}^{-1} \bar{\Phi}_s \tau_{11} \\ & + \tau_{31}^T \tau_{11}^T \bar{\Phi}_s^T \bar{R}_{11}^{-1} \bar{\Phi}_s \tau_{11} \tau_{31}, \end{aligned} \quad (5.15)$$

$$\begin{aligned} 0 = & \bar{A}_3 Q + Q \bar{A}_3^T + \tau_{21} \bar{Q}_s \bar{V}_{12}^{-1} \bar{Q}_s^T \tau_{21}^T \\ & - \tau_{31} \tau_{21} \bar{Q}_s \bar{V}_{12}^{-1} \bar{Q}_s^T \tau_{21}^T \tau_{31}^T - G^T K_3 C_2 Q - Q(K_3 C_2)^T G, \end{aligned} \quad (5.16)$$

$$\begin{aligned} 0 = & \bar{A}_4^T \bar{P} + \bar{P} \bar{A}_4 + \tau_{11}^T \bar{\Phi}_s^T \bar{R}_{11}^{-1} \bar{\Phi}_s \tau_{11} \\ & - \tau_{31}^T \tau_{11}^T \bar{\Phi}_s^T \bar{R}_{11}^{-1} \bar{\Phi}_s \tau_{11} \tau_{31} - \Gamma^T K_2 B_2 P - P(K_2 B_2)^T \Gamma, \end{aligned} \quad (5.17)$$

where

$$\bar{A}_1 = \bar{A} + B_2 K_1 C_2 - G^T K_3 C_2, \quad \bar{A}_2 = \bar{A} + B_2 K_1 C_2 + B_2 K_2 \Gamma,$$

$$\bar{A}_3 = \bar{A}_4 = \bar{A}_0 + B_2 K_1 C_2 - G^T K_3 C_2, \quad \bar{A}_5 = \bar{A}_0 + B_2 K_1 C_2 + B_2 K_2 \Gamma.$$

VI. DIRECTIONS FOR FURTHER RESEARCH

More general solutions can be obtained by incorporating singular estimation techniques [15] where noise-free measurements are repeatedly differentiated to enlarge the class of available outputs.

ACKNOWLEDGMENT

The author wishes to thank D. C. Hyland for numerous helpful discussions.

REFERENCES

- W. S. Levine and M. Athans, "On the determination of the optimal constant output feedback gains for linear multivariable systems," *IEEE Trans. Automat. Contr.*, vol. AC-15, pp. 44–48, 1970.
- J. Medanic, "On stabilization and optimization by output feedback," *Proc. 12th Asilomar Conf. Circ., Syst. Comp.*, Pacific Grove, CA, Nov. 1978, pp. 412–416.
- S. Basuthakur and C. H. Knapp, "Optimal constant controllers for stochastic linear systems," *IEEE Trans. Automat. Contr.*, vol. AC-20, pp. 664–666, 1975.
- D. C. Hyland and D. S. Bernstein, "The optimal projection equations for fixed-order dynamic compensation," *IEEE Trans. Automat. Contr.*, vol. AC-29, pp. 1034–1037, 1984.
- W. M. Wonham, "Optimal stationary control of a linear system with state-dependent noise," *SIAM J. Contr.*, vol. 5, pp. 486–500, 1967.
- D. Kleinman, "Optimal stationary control of linear systems with control-dependent noise," *IEEE Trans. Automat. Contr.*, vol. AC-14, pp. 673–677, 1969.
- P. McLane, "Optimal stochastic control of linear systems with state and control-dependent disturbances," *IEEE Trans. Automat. Contr.*, vol. AC-16, pp. 793–798, 1971.
- J. Bismut, "Linear-quadratic optimal stochastic control with random coefficients," *SIAM J. Contr.*, vol. 14, pp. 419–444, 1976.
- G. N. Mil'stein, "Design of stabilizing controller with incomplete state data for linear stochastic system with multiplicative noise," *Automat. Remote Contr.*, vol. 43, pp. 653–659, 1982.
- Y. A. Phillis, "Controller design of systems with multiplicative noise," *IEEE Trans. Automat. Contr.*, vol. AC-30, pp. 1017–1019, 1985.
- D. S. Bernstein and D. C. Hyland, "The optimal projection/maximum entropy approach to designing low-order, robust controllers for flexible structures," in *Proc. 24th IEEE Conf. Decision Contr.*, Fort Lauderdale, FL, Dec. 1985, pp. 745–752.
- D. S. Bernstein and S. W. Greeley, "Robust controller synthesis using the maximum entropy design equations," *IEEE Trans. Automat. Contr.*, vol. AC-31, pp. 362–364, 1986.
- S. L. Campbell and C. D. Meyer, Jr., *Generalized Inverses of Linear Transformations*. London: Pitman, 1979.
- C. R. Rao and S. K. Mitra, *Generalized Inverse of Matrices and Its Applications*. New York: Wiley, 1971.
- V. B. Haas, "Reduced order state estimation for linear systems with exact measurements," *Automatica*, vol. 20, pp. 225–229, 1984.

**A Riccati Equation Approach
to the
Singular LQG Problem**

(ABSTRACT)

by

**Yoram Halevi
Faculty of Mech. Engineering
Institute of Technology
Haifa 32000
Israel**

**Wassim M. Haddad
Department of Mechanical
and Aerospace Engineering
Florida Institute of Technology
Melbourne, FL 32901**

**Dennis S. Bernstein
Government Aerospace
Systems Division
Harris Corporation,
Melbourne, FL 32902**

1. Introduction

The singular LQG control problem has been of considerable interest for almost two decades ([1-15]). Such problems arise when some of the measurements are noise free or when some of the control signals are unweighted. This will be the case, for example, if the sensor noise is colored or if actuator dynamics are included. Augmentation of the plant dynamics by means of noise filters or actuator dynamics thus leads directly to the singular problem formulation.

Most of the literature on the singular LQG problem is based upon limiting procedures in which suitable weighting matrices and noise intensities approach zero. These results demonstrate the types of behavior which can arise in the limiting solution including impulsive controls and singular arcs.

The available literature is concerned, of course, with determining the optimal limiting (i.e., singular) control. In practical applications, however, it is often of interest to determine the optimal controller within a prespecified class of controllers. In particular, we consider the singular LQG problem in which the controller is *preconstrained* to possess a fixed dynamic feedback structure. One benefit of this approach is that the fixed structure constraint eliminates the possibility of impulsive controls and other complex behavior.

Preliminary results for the singular LQG problem were obtained in [15] using the fixed structure approach. For generality, the problem considered in [15] permits the design of fixed-order, i.e., reduced-order, dynamic compensators. As in [16], the solution is given by a system of coupled algebraic Riccati and Lyapunov equations whose solutions (denoted by Q, P, \hat{Q}, \hat{P}) are used to explicitly characterize the optimal feedback gains. The coupling is due to a pair of oblique projections (i.e., idempotent matrices) which arise as a direct consequence of the fixed structure constraint. The order-reduction projection τ defined by

$$\tau \triangleq \hat{Q} \hat{P} (\hat{Q} \hat{P})^\#,$$

where $()^\#$ denotes group generalized inverse, appeared originally in [16], while the static projection ν given by

$$\nu \triangleq Q C^T (C Q C^T)^{-1} C,$$

is familiar from least squares analysis.

The results of [15] are incomplete, however, in that the gains associated with certain feedback paths were not given explicitly. For the corresponding singular estimation problem ([17]) this de-

fect was remedied in [18] where all feedback gains were explicitly characterized. In addition, the solution obtained in [18] was shown to agree completely with results obtained using standard limiting methods when the (unconstrained) optimal singular estimator does not possess differentiators ([19]). The results of [18] thus provide an alternative approach to the singular estimation problem considered in [20,21,22] and the numerous references therein.

The contribution of the present paper is thus to complete the development of [15] by incorporating the methods used in [18]. Accordingly, we derive a coupled system of modified Riccati and Lyapunov equations which explicitly characterize the feedback gains of the fixed-structure singular LQG controller. For generality we consider partial or total singularity in both the control weighting and measurement noise intensity matrices, and we allow the dynamic compensator to be of arbitrary dimension less than or equal to the number of plant states minus the number of noise-free measurements. In the special case in which the order of the dynamic compensator is *equal* to the number of plant states minus the number of noise-free measurements (i.e., the quasi full-order case), then we show that the optimal solution decomposes (separates) into a reduced-order observer followed by state feedback.

An additional benefit of our approach is the ability to impose an upper bound on the number of differentiators to be included in the feedback controller. That is, while certain measurement signals may be noise free and hence differentiable, it may be undesirable in practice to implement more than one level of differentiation or, perhaps any differentiation at all. Furthermore, as in [18] we demonstrate connections with earlier results by showing that the fixed structure solution agrees with the standard limiting solution when the latter possesses the same number of differentiators as are included in the prespecified controller structure.

To illustrate the solution we consider a numerical example of fourth order with two noise-free measurements and one noisy measurement. (Numerical results for the singular control case are immediate from duality). By solving the coupled systems of modified Riccati and Lyapunov equations by means of a homotopy algorithm ([23]), we obtain the quasi full-order solution (second-order controller) as well as an optimal first-order controller.

References

1. B. Friedland, "Limiting Forms of Optimum Stochastic Linear Regulators," *Trans. ASME J. Dyn. Sys. Meas. Contr.*, Vol. 93, pp. 134-141, 1971.
2. Y.-C. Ho, "Linear Stochastic Singular Control Problems," *J. Optim. Thy. Appl.*, Vol. 9, pp. 24-31, 1972.
3. M. F. Hutton, "Solutions of the Singular Stochastic Regulator Problems," *Trans. ASME J. Dyn. Sys. Meas. Contr.*, Vol. 95, pp. 414-417, 1973.
4. H. G. Kwatny, "Minimal Order Observers and Certain Singular Problems of Optimal Estimation and Control," *IEEE Trans. Autom. Contr.*, Vol. AC-19, pp. 274-276, 1974.
5. R. E. O'Malley, Jr., and A. Jameson, "Singular Perturbations and Singular Arcs - Part I," *IEEE Trans. Autom. Contr.*, Vol. AC-20, pp. 218-226, 1975.
6. R. E. O'Malley, Jr., and A. Jameson, "Singular Perturbations and Singular Arcs - Part II," *IEEE Trans. Autom. Contr.*, Vol. AC-22, pp. 328-337, 1977.
7. B. A. Francis and K. Glover, "Bounded Peaking in the Optimal Linear Regulator With Cheap Control," *IEEE Trans. Autom. Contr.*, Vol. AC-23, pp. 608-617, 1978.
8. B. A. Francis, "The Optimal Linear-Quadratic Time-Invariant Regulator with Cheap Control," *IEEE Trans. Autom. Contr.*, Vol. AC-24, pp. 616-621, 1979.
9. V. Dragan and A. Halanay, "Cheap Control and Singularly Perturbed Matrix Riccati Differential Equations," *Rev. Roum. Math. Pures et. Appl.*, Vol. 26, pp. 21-40, 1981.
10. V. B. Haas, "The Singular Steady State Linear Regulator," *SIAM J. Contr. Optim.*, Vol. 20, pp. 247-257, 1982.
11. J. O'Reilly, "Partial Cheap Control of the Time-Invariant Regulator," *Int. J. Contr.*, Vol. 37, pp. 909-927, 1983.
12. P. Sannuti, "Direct Singular Perturbation Analysis of High-Gain and Cheap Control Problems," *Automatica*, Vol. 19, pp. 41-51, 1983.
13. M. L. J. Hautus and L. M. Silverman, "System Structure and Singular Control," *Lin. Alg. Appl.*, Vol. 50, pp. 369-402, 1983.
14. Y. Halevi and Z. J. Palmor, "Extended Limiting Forms of Optimum Observers and LQG Regulators," Vol. 43, pp. 193-212, 1986.
15. D. S. Bernstein, "The Optimal Projection for Static and Dynamic Output Feedback: The Singular Case," *IEEE Trans. Autom. Contr.*, Vol. AC-32, pp. 1139-1143, 1987.
16. D. C. Hyland and D. S. Bernstein, "The Optimal Projection Equations for Fixed-Order Dynamic Compensation," *IEEE Trans. Autom. Contr.*, Vol. AC-29, pp. 1034-1037, 1984.
17. W. M. Haddad and D. S. Bernstein, "The Optimal Projection Equations for Reduced Order State Estimation: The Singular Measurement Noise Case," *IEEE Trans. Autom. Contr.*, Vol. AC-32, pp. 1135-1139, 1987.

18. Y. Halevi, "The Optimal Reduced Order Estimator for Systems with Singular Measurement Noise," *IEEE Trans. Autom. Contr.*, Vol. 34, 1989.
19. Y. Halevi, "Limiting Forms of Optimal Observers," *Int. J. Contr.*, Vol. 47, pp. 209-216, 1988.
20. A. E. Bryson and D. E. Johansen, "Linear Filtering for Time Varying Systems Using Measurements Containing Colored Noise," *IEEE Trans. Autom. Contr.*, Vol. AC-10, pp. 4-10, 1965.
21. J. M. Schumacher, "A Geometric Approach to the Singular Filtering Problem," *IEEE Trans. Autom. Contr.*, AC-30, pp. 773-776, 1985.
22. E. Soroka and U. Shaked, "Direct Solution to the General Reduced Order Stochastic Observation Problem," *Int. J. Contr.*, Vol. 45, pp. 713-728, 1987.
23. S. Richter, "A Homotopy Algorithm for Solving the Optimal Projection Equations for Fixed-Order Dynamic Compensation: Existence, Convergence, and Global Optimality," *Proc. Amer. Contr. Conf.*, pp. 1527-1531, Minneapolis, MN, June 1987.

APPENDIX F: Stochastic Modeling

- 35. D. S. Bernstein and D. C. Hyland, "The Optimal Projection/Maximum Entropy Approach to Designing Low-Order, Robust Controllers for Flexible Structures," Proc. 24th IEEE Conf. Dec. Contr., pp. 745-752, Fort Lauderdale, FL, December 1985.
- 36. D. S. Bernstein, L. D. Davis, S. W. Greeley and D. C. Hyland, "Numerical Solution of the Optimal Projection/Maximum Entropy Design Equations for Low-Order, Robust Controller Design," Proc. 24th IEEE Conf. Dec. Contr., pp. 1795-1798, Fort Lauderdale, FL, December 1985.
- 104. D. S. Bernstein and D. C. Hyland, "Optimal Projection Equations for Reduced-Order Modelling, Estimation and Control of Linear Systems with Multiplicative White Noise," J. Optim. Thy. Appl., Vol. 58, pp. 387-409, 1988.
- 39. D. S. Bernstein and S. W. Greeley, "Robust Controller Synthesis Using the Maximum Entropy Design Equations," IEEE Trans. Autom. Contr., Vol. AC-31, pp. 362-364, 1986.
- 77. D. S. Bernstein, "Robust Static and Dynamic Output-Feedback Stabilization: Deterministic and Stochastic Perspectives," IEEE Trans. Autom. Contr., Vol. AC-32, pp. 1076-1084, 1987.

THE OPTIMAL PROJECTION/MAXIMUM ENTROPY APPROACH TO DESIGNING LOW-ORDER, ROBUST CONTROLLERS FOR FLEXIBLE STRUCTURES

Dennis S. Bernstein and David C. Hyland

Harris Corporation
Government Aerospace Systems Division
Controls Analysis and Synthesis Group
Melbourne, Florida 32902

Abstract. The Optimal Projection/Maximum Entropy approach to designing low-order controllers for high-order systems with parameter uncertainties is reviewed.

The philosophy of representing uncertain parameters by means of Stratonovich multiplicative white noise is motivated by means of the Maximum Entropy Principle of Jaynes and statistical analysis of modal systems. The main result, the optimal projection equations for fixed-order dynamic compensation in the presence of state-, control- and measurement-dependent noise, represents a fundamental generalization of classical LQG theory.

1. Overview

Optimal Projection/Maximum Entropy Stochastic Modelling and Reduced-Order Design Synthesis is a rigorous new approach to designing robust, implementable feedback controllers. Inspired by Statistical Energy Analysis [1], a branch of dynamic modal analysis developed for analyzing acoustic vibrations, its present stage of development [2-22], embodies a mathematically rigorous, fundamental generalization of classical steady-state Kalman filter and linear-quadratic-Gaussian (LQG) optimal control theory. Although LQG theory is an effective tool for optimally quantifying performance/sensor-resolution and performance/actuation-level tradeoffs, it suffers from two fundamental defects which severely limit its usefulness in practice.

1. Whereas the dimension of an LQG controller must equal that of the controlled plant, optimal projection design characterizes the quadratically optimal controller of fixed dimension less than that of the plant in accordance with implementation constraints (e.g., reliability, complexity or real-time computing capability).

2. Whereas LQG presumes exact knowledge of each and every parameter appearing in the state-space plant description, maximum entropy modelling provides a stochastic plant model which admits ignorance with regard to parameter values in accordance with unavoidable plant modelling errors.

with regard to the latter item, it should be stressed that one of the major problems in designing high-performance control systems is that of robustness, i.e., the ability of the controller to tolerate errors in the plant model upon which its design is predicated. Maximum entropy modelling directly addresses this problem by incorporating into the dynamic model a representation of ignorance (i.e., uncertainty) regarding physical parameters. Roughly speaking, the idea behind the approach is to use a probabilistic representation of each imperfectly known plant parameter so that the quadratically optimal control system designed under this probabilistic model is automatically desensitized to actual parameter variations when the control system is implemented. The overall control-design procedure thus avoids laborious trial and error post-design "tweaking."

2. Motivation

The inherent time- and frequency-domain duality in representing linear dynamic systems (i.e., state space versus transfer functions) provides control-system designers with complementary methodologies for assessing tradeoffs between performance objectives and the design constraints of sensor resolution, actuation levels, plant modelling accuracy and controller complexity. In spite of the ability of LQG to optimally quantify performance/sensor-resolution and performance/actuation-level tradeoffs in a state-space setting, its enormous sensitivity to plant modelling errors has forced practitioners to seek generalizations of classical frequency domain methods. In numerous practical situations, however, input/output techniques possess fundamental limitations. For example, representing modelling uncertainty in a frequency-domain plant model $G(s)$ by means of

$$G(s) + \Delta G(s),$$

where ΔG remains in a normed neighborhood of G , is essentially a black-box (nonparametric) approach: By failing to exploit physical laws (such as conservation of energy), systems represented by $G + \Delta G$ may actually be physically impossible, resulting in unwarranted design conservatism at the expense of system performance. Hence, when some knowledge of internal mechanisms is available (i.e., the "grey-box" situation), state-space representations may provide greater modelling fidelity. These observations are motivated by the problem of controlling vibration in flexible structures

where internal energy dissipation precludes right-half-plane poles and where high-order finite-element models have highly structured dynamics but possess numerous uncertain parameters. More specifically, frequency uncertainties for higher order modes are much larger in magnitude than damping uncertainties. Hence, the inability to differentiate between these physical parameters in an input-output representation leads to severe performance consequences.

The optimal projection/maximum entropy approach generalizes LQG theory in two fundamental respects: design of reduced-order controllers plus accommodation of a priori parameter uncertainties. For clarity, we discuss these generalizations separately following the left branch of Fig. 1. Optimal projection design is discussed in Section 3, followed by maximum entropy modelling in Sections 4, 5, and 6.

OPTIMAL PROJECTION/MAXIMUM ENTROPY APPROACH TO LOW-ORDER, ROBUST CONTROLLER DESIGN

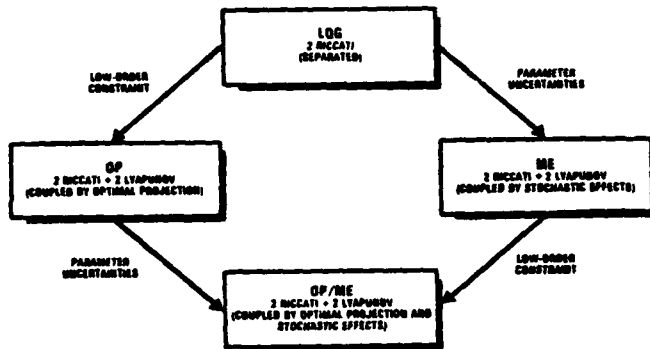


Fig. 1

3. Review of the Optimal Projection Approach

Most research into the design of reduced-order controllers involves one of two sequential procedures: model reduction followed by controller design, or controller design followed by controller reduction. The optimal projection equations represent a radical departure from both of these approaches by directly characterizing the quadratically optimal reduced-order controller for a high-order model. Assuming a purely dynamic linear structure for the desired compensator, whose order is determined by implementation constraints, a parameter optimization approach is taken. There is, of course, nothing novel about this approach per se and it has been widely studied in the control literature (see references listed in [18]). This approach, however, fell into disrepute because of the extreme complexity of the grossly unwieldy first-order necessary conditions which afforded little insight and engendered brute-force gradient search techniques. The crucial discovery occurred [7] where it was revealed that the necessary conditions for the dynamic-compensation problem give rise to the definition

of an optimal projection as a rigorous, unassailable consequence of quadratic optimality without recourse to ad hoc methods. Exploitation of this projection leads to immense simplification of the "primitive" form of the necessary conditions which now provide a transparent generalization of the pair of separated Riccati equations of standard LQG theory. In particular, the optimal projection equations comprise a system of four matrix equations coupled by an oblique projection which determines the optimal controller gains. The system of matrix equations includes a pair of modified Riccati equations which are analogous to the standard Riccati equations, along with a pair of modified Lyapunov equations which arise separately in the model reduction problem [19]. The coupling by means of the projection reveals the inherent inseparability of these operations in the reduced-order case since optimality considerations demand that, in a very precise sense, "reduction" and "control design" be performed simultaneously. Hence the full-order model is retained throughout the control design process and there is no need to truncate the plant model.

4. Maximum Entropy Modelling

Although optimal projection design deals directly and rigorously with the question of system dimension by trading order off against performance, it is, nevertheless, predicated upon the availability of a completely accurate plant and disturbance model. Maximum entropy modelling, however, addresses the robustness problem by directly including parameter uncertainties in the plant and disturbance models so that optimal projection design plus maximum entropy modelling automatically yields control designs that trade performance off against modelling uncertainties. In order to review the maximum entropy approach it is important to discuss the class of problems that motivated this work, namely, control of flexible structures. A finite-element model of a large flexible structure is, generally, an extremely high-order system. For example, a version of the widely studied Draper Model #2 includes 3 rigid body modes, 147 elastic modes and 6 disturbance states, i.e., a total of 306 states, along with 9 sensors and 9 actuators. Besides the high order of these systems, finite element modelling is known to have poor accuracy, particularly for the high-order modes. Reasonable and not overly conservative uncertainty estimates predict 30-50 percent error in modal frequencies after the first 10 modes, with the situation considerably more complex (and pessimistic) for damping estimates.

Maximum entropy modelling is a form of stochastic modelling. Although external disturbances are traditionally modelled as random processes, the use of stochastic theory to model plant parameter uncertainty has seen relatively limited application. To dispel all objections to a stochastic parameter-uncertainty model, we invoke the modern information-theoretic interpretation of probability theory. Rather than regard the probability of an event as an objective quantity such as the limiting frequency of outcomes of numerous repetitions (as, e.g., the number of heads in 1,000 coin tosses), we adopt the view that the probability of an event is a subjective quantity which reflects the observer's certainty as to

whether a particular event will or will not occur. This quantity is nothing more than a measure of the information (including, e.g., all theoretical analysis and empirical data) available to the observer. In this sense the validity of a stochastic model of a flexible structure, for example, does not rely upon the existence of a fleet of such objects (substitute "ensemble" for "fleet" in the classical terminology) but rather resides in the interpretation that it expresses the engineer's certainty or uncertainty regarding the values of physical parameters such as stiffnesses of structural components. This view of probability theory has its roots in Shannon's information theory but was first articulated unambiguously by Jaynes [23-26].

The preeminent problem in modelling the real world is thus the following: Given limited (incomplete) a priori data, how does one construct a well-defined (complete) probability model which is consistent with the available data but which avoids inventing data which does not exist? To this end we invoke Jaynes' Maximum Entropy Principle: First, define a measure of ignorance in terms of the information-theoretic entropy, and then determine the probability distribution which maximizes this measure subject to agreement with the available data. The reasoning behind this principle is that the probability distribution which maximizes the a priori ignorance must be the least presumptive (i.e., least likely to invent data) on the average since the corresponding amount of a posteriori learned information (should all uncertainty suddenly disappear) would necessarily be maximized. If, for some probability distribution, the a priori ignorance and hence the a posteriori learning were less than their potentially maximum value, then this distribution must be based upon invented and hence generally incorrect data. The Maximum Entropy Principle is clearly desirable for control-system design where the introduction of false data is to be assiduously avoided.

5. Minimum-Information Modelling of Parameter Uncertainties

For dynamical-system modelling it was first shown by Hyland [2] that for structural systems the minimum information linear stochastic dynamic model induced by the Maximum Entropy Principle of Jaynes is a Stratonovich multiplicative white noise model. In the present paper we adopt this model and explore its ramifications for general systems. The basic model is given by

$$\dot{\tilde{x}}(t) = (\tilde{A} + \sum_{i=1}^p v_i(t) \tilde{A}_i) \tilde{x}(t) + \tilde{w}(t), \quad (5.1)$$

where $\tilde{x}(t) \in \mathbb{R}^n$, \tilde{A} , $\tilde{A}_i \in \mathbb{R}^{n \times n}$, $\tilde{w}(t)$ is zero-mean Gaussian white disturbance noise with non-negative-definite intensity \tilde{V} , and $v_i(t)$ are zero-mean, unit-intensity Gaussian white noise processes which are mutually uncorrelated and uncorrelated with $\tilde{w}(t)$. The multiplicative white noise model (5.1) can be regarded as a parameter uncertainty model where each $v_i(t)$ corresponds to a single uncertain parameter whose pattern and magnitude are given by $\tilde{A}_i / \|\tilde{A}_i\|$ and $\|\tilde{A}_i\|$, respectively.

To see why (5.1) is a minimum information model of parameter uncertainty, note that when the pattern $\tilde{A}_i / \|\tilde{A}_i\|$ of an uncertain parameter

is known, all available data (theoretical and empirical) can be used to determine a suitable value for the magnitude $\|\tilde{A}_i\|$ to reflect the

corresponding level of uncertainty. Clearly, the collection of magnitudes constitutes the minimum data set needed to render (5.1) well defined. For the harmonic oscillator with uncertain natural frequency, the uncertainty magnitude is given by the reciprocal of the decorrelation time (Fig. 2). Note that the uncertainty representation (5.1) is a minimum information model in the sense that it eschews detailed descriptions of joint probability statistics of unknown parameters.

MINIMUM-INFORMATION MODELLING DECORRELATION TIME

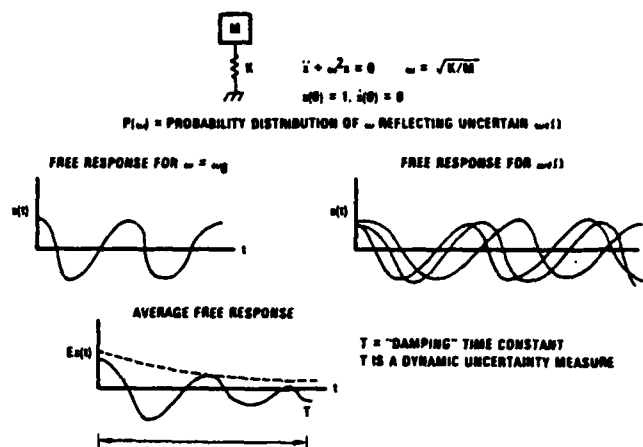


Fig. 2

To eliminate the white noise formalism, the model (5.1) is usually rigorized by the Ito differential equation

$$d\tilde{x}_t = (\tilde{A}dt + \sum_{i=1}^p dv_{it} \tilde{A}_i) \tilde{x}_t + d\tilde{w}_t, \quad (5.2)$$

where dv_{it} and $d\tilde{w}_t$ are Brownian motions, i.e., Wiener processes. Although such models were studied extensively for control design, this approach fell into disrepute with the publication of [27, 28] where it was shown for discrete-time systems that sufficiently high uncertainty levels (i.e., magnitudes $\|\tilde{A}_i\|$ above a "threshold")

lead to the nonexistence of a steady-state solution. Although it was purported that this "phenomenon" was an "obvious" consequence of high uncertainty levels, these conclusions failed to take into account (possibly because of the discrete-time setting) the subtle relationship between the ordinary differential equation (5.1) and the stochastic differential equation (5.2). Indeed, it was shown in [29] that if a stochastic differential equation is regarded as the limit of a sequence of approximating ordinary differential equations, then (5.2) is not the correct version of (5.1).

Instead, the ordinary differential equation (5.1) with multiplicative white noise corresponds to the corrected Ito differential equation

$$d\tilde{x}_t = (\tilde{A}_s dt + \sum_{i=1}^p dv_{it} \tilde{A}_i) \tilde{x}_t + d\tilde{w}_t, \quad (5.3)$$

where

$$\tilde{A}_s = \tilde{A} + \frac{1}{2} \sum_{i=1}^p \tilde{A}_i^2, \quad (5.4)$$

which differs from the "naive" equation (5.2) by a systematic drift term. The form of (5.3) was corroborated completely independently by Stratonovich [30], whose results actually appeared in the Russian literature prior to 1965. His approach is based upon an alternative definition of the stochastic integral which differs from the Ito definition by a mathematical technicality.

In spite of the glaring technicality of the Stratonovich correction, almost all research on the estimation and control of such systems failed to perceive its physical significance. Specifically, the Stratonovich correction neutralizes the "threshold uncertainty principle": For systems which are inherently stable under particular parameter variations (e.g., structures with uncertain stiffness matrices), the Stratonovich formulation correctly predicts unconditional second-moment stability in contrast to the Ito formulation within which a stringent uncertainty threshold is encountered. We shall now proceed to demonstrate this fact by means of a compelling example relevant to the modeling and control of flexible structures, in particular, and hyperbolic systems, in general.

First, suppose that zero-point deviations of $\tilde{x}(t)$ are of interest and are evaluated according to

$$J = \lim_{t \rightarrow \infty} E \tilde{x}(t)^T \tilde{R} \tilde{x}(t) = \lim_{t \rightarrow \infty} \text{tr } \tilde{Q}(t) \tilde{R}, \quad (5.5)$$

where $\tilde{R} \in \mathbb{R}^{n \times n}$ and the second moment of the state is

$$\tilde{Q}(t) \triangleq E[\tilde{x}(t) \tilde{x}(t)^T]. \quad (5.6)$$

The obvious fact cannot be overemphasized that the primary state statistic of design interest in linear-quadratic optimization is the state covariance (5.6). From Ito calculus it follows that $\tilde{Q}(t)$ is given for the naive model (5.2) by

$$\dot{\tilde{Q}}(t) = \tilde{A} \tilde{Q}(t) + \tilde{Q}(t) \tilde{A}^T + \sum_{i=1}^p \tilde{A}_i \tilde{Q}(t) \tilde{A}_i^T + \tilde{V} \quad (5.7)$$

and for the corrected model (5.3) by

$$\dot{\tilde{Q}}(t) = \tilde{A}_s \tilde{Q}(t) + \tilde{Q}(t) \tilde{A}_s^T + \sum_{i=1}^p \tilde{A}_i \tilde{Q}(t) \tilde{A}_i^T + \tilde{V}. \quad (5.8)$$

Each of these "stochastic" Lyapunov differential equations, which govern the evolution of the second moment, should be regarded as $n(n+1)/2$ ordinary differential equations. Hence we wish

to address the following question: How do the solutions of the stochastic Lyapunov equations (5.7) and (5.8) differ from each other and from the "deterministic" Lyapunov equation

$$\dot{\tilde{Q}}(t) = \tilde{A} \tilde{Q}(t) + \tilde{Q}(t) \tilde{A}^T + \tilde{V}, \quad (5.9)$$

particularly in the presence of high uncertainty levels? The answer to this question of course depends upon the stochastic modification terms which for the naive model are given by

$$M_1[\tilde{Q}(t)] \triangleq \sum_{i=1}^p \tilde{A}_i \tilde{Q}(t) \tilde{A}_i^T \quad (5.10)$$

and for the corrected model by

$$M_2[\tilde{Q}(t)] = \sum_{i=1}^p [\frac{1}{2} \tilde{A}_i^2 \tilde{Q}(t) + \frac{1}{2} \tilde{Q}(t) \tilde{A}_i^2 + \tilde{A}_i \tilde{Q}(t) \tilde{A}_i^T]. \quad (5.11)$$

Consider a system consisting of a pair of lightly damped modes so that

$$\tilde{A} \triangleq \begin{bmatrix} 0 & -\omega_1 & 0 & 0 \\ \omega_1 & -2\eta_1 & 0 & 0 \\ 0 & 0 & 0 & -\omega_2 \\ 0 & 0 & \omega_2 & -2\eta_2 \end{bmatrix},$$

where $\eta_1 = \zeta_1 \omega_1$, and to represent frequency uncertainties let

$$\tilde{A}_1 = \gamma_1 \begin{bmatrix} 0 & -1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix},$$

$$\tilde{A}_2 = \gamma_2 \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 \end{bmatrix},$$

where for simplicity we have ignored the effects of frequency uncertainties on the effective decay rate η_i . The magnitudes of the uncertainties are scaled by means of γ_1 and γ_2 . For this example the Ito stochastic modification

$M_1[\tilde{Q}(t)]$ has the form

$$\begin{bmatrix} \gamma_1^2 \tilde{Q}_{22}(t) & -\gamma_1^2 \tilde{Q}_{12}(t) & 0 & 0 \\ -\gamma_1^2 \tilde{Q}_{12}(t) & \gamma_1^2 \tilde{Q}_{11}(t) & 0 & 0 \\ 0 & 0 & \gamma_2^2 \tilde{Q}_{44}(t) & -\gamma_2^2 \tilde{Q}_{34}(t) \\ 0 & 0 & -\gamma_2^2 \tilde{Q}_{34}(t) & \gamma_2^2 \tilde{Q}_{33}(t) \end{bmatrix}$$

Although the off-diagonal terms have a stabilizing effect, it is clear that the diagonal elements destabilize the state variances. Hence, it is not surprising that for sufficiently high uncertainty levels, i.e., $\gamma_i \gg 0$, the Ito model is second-moment unstable. These observations are completely in accordance with the threshold uncertainty principle. The Stratonovich corrected stochastic modification $M_S[\dot{Q}(t)]$, however, has the form

$$\begin{bmatrix} \gamma_1^2[\tilde{u}_{22}(t) - \tilde{u}_{11}(t)] & -2\gamma_1^2\tilde{u}_{12}(t) & -\frac{1}{2}(\gamma_1^2\gamma_2^2)\tilde{u}_{13}(t) & -\frac{1}{2}(\gamma_1^2\gamma_2^2)\tilde{u}_{14}(t) \\ -2\gamma_1^2\tilde{u}_{12}(t) & \gamma_1^2[\tilde{u}_{11}(t) - \tilde{u}_{22}(t)] & -\frac{1}{2}(\gamma_1^2\gamma_2^2)\tilde{u}_{23}(t) & -\frac{1}{2}(\gamma_1^2\gamma_2^2)\tilde{u}_{24}(t) \\ -\frac{1}{2}(\gamma_1^2\gamma_2^2)\tilde{u}_{13}(t) & -\frac{1}{2}(\gamma_1^2\gamma_2^2)\tilde{u}_{32}(t) & \gamma_2^2[\tilde{u}_{44}(t) - \tilde{u}_{33}(t)] & -2\gamma_2^2\tilde{u}_{34}(t) \\ -\frac{1}{2}(\gamma_1^2\gamma_2^2)\tilde{u}_{14}(t) & -\frac{1}{2}(\gamma_1^2\gamma_2^2)\tilde{u}_{42}(t) & -2\gamma_2^2\tilde{u}_{34}(t) & \gamma_2^2[\tilde{u}_{33}(t) - \tilde{u}_{44}(t)] \end{bmatrix}$$

which also has stabilizing off-diagonal elements but has fundamentally different diagonal elements: Rather than destabilizing the state variances, the diagonal elements of the corrected stochastic modification are equilibrating. This effect is even more striking when M_1 and M_5 are transformed into the basis with respect to which

$$\tilde{A} = \begin{bmatrix} -j\omega_1 - \eta_1 & 0 & 0 & 0 \\ 0 & j\omega_1 - \eta_1 & 0 & 0 \\ 0 & 0 & -j\omega_2 - \eta_2 & 0 \\ 0 & 0 & 0 & j\omega_2 - \eta_2 \end{bmatrix}$$

where higher order terms in η have been ignored. In this basis, the diagonal terms of $M_1[\dot{Q}(t)]$ are destabilizing whereas the diagonal terms of $M_S[\dot{Q}(t)]$ exactly vanish.

The negative coefficients in the off-diagonal terms imply progressive decorrelation between pairs of dynamical states. This informational or statistical damping phenomenon is a direct result of parameter uncertainties captured by the multiplicative white noise model. The Stratonovich correction, moreover, is crucial: By neutralizing the threshold uncertainty principle, it permits the consideration of long-term effects for arbitrary uncertainty levels.

As an example of the ramifications of these observations, assume (as is usually the case in practice) that uncertainties in modal frequency obtained from finite-element analysis of a flexible structure increase with mode number. From the form of $M_S[\dot{Q}(t)]$ it is easy to deduce that the steady-state covariance

$$\lim_{t \rightarrow \infty} \dot{q}(t)$$

satisfying

$$0 = \tilde{\chi}_s \tilde{Q} + \tilde{Q} \tilde{A}_s^T + \sum_{i=1}^p \tilde{\chi}_i \tilde{Q} \tilde{A}_i^T + \tilde{V} \quad (5.12)$$

becomes increasingly diagonally dominant with increasing frequency and thus assumes the qualitative form given in Fig. 3. The benefits of this sparse form are important: The computational effort required to determine the steady-state covariance (and thus to design a closed-loop controller, for example) is directly proportional to the amount of information reposed in the model or, equivalently, inversely proportional to the level of modelled parameter uncertainty. This casts new light on the computational design burden vis-a-vis the modelling question: The computational burden depends only upon the information actually available. A simple control-design exercise involving full-state feedback illustrates this point. The gains for the higher order modes of the beam in Fig. 4, whose frequency uncertainties increase linearly with frequency, were obtained with modest computational effort in spite of $\tilde{n} = 100$ (see Fig. 5). Another important ramification of the qualitative form of \tilde{Q} is the automatic generation of a high-/low-authority control law. Note that for the higher order and hence highly uncertain modes the control gains indicate an inherently stable, low-performance rate-feedback control law, whereas for the lowest-order modes the control law is high authority, i.e., "LQ" in character.

EFFECT OF FREQUENCY UNCERTAINTIES ON THE QUALITATIVE STRUCTURE OF THE STEADY-STATE COVARIANCE $Q = \lim_{t \rightarrow \infty} E[x(t)x(t)^T]$

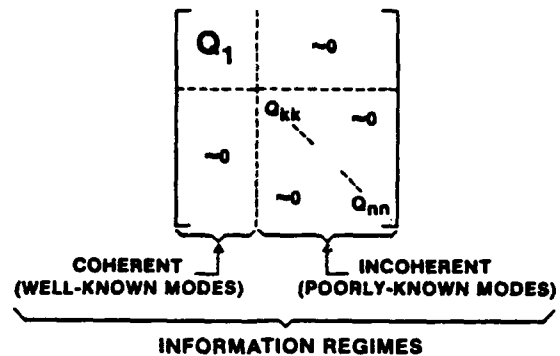


Fig. 3

FULL-STATE FEEDBACK CONTROL SCHEME

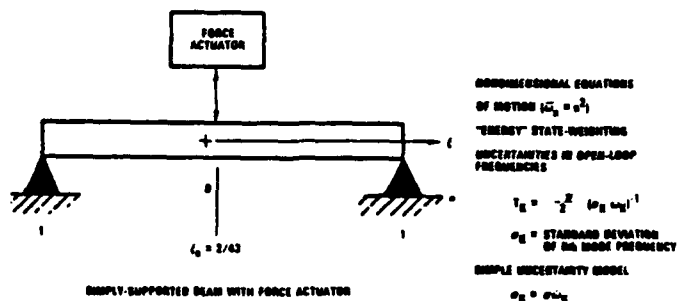


Fig. 4

STOCHASTIC BEAM PROBLEM—GAIN MAGNITUDES

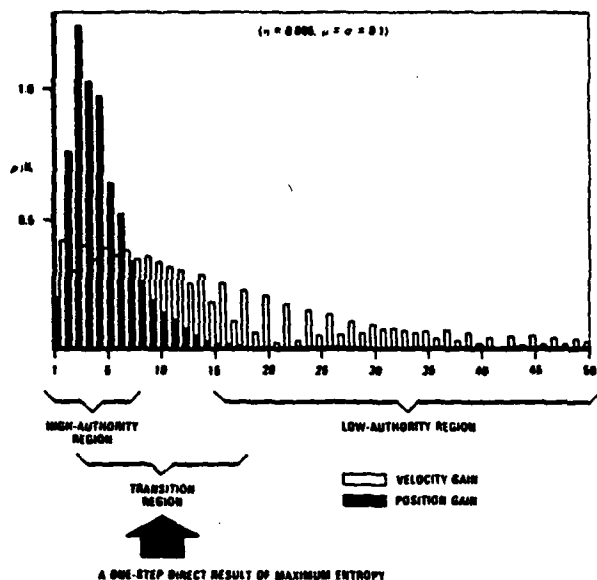


Fig. 5

6. Optimal Projection Design with Stratonovich Multiplicative White Noise

To state the Reduced-Order Dynamic Compensation Problem, we require the following notation. Let

$$\begin{aligned} x \in \mathbb{R}^n, y \in \mathbb{R}^l, u \in \mathbb{R}^m, A, A_1, \dots, A_p \in \mathbb{R}^{n \times n}, \\ B, B_1, \dots, B_p \in \mathbb{R}^{n \times m}, C, C_1, \dots, C_p \in \mathbb{R}^{l \times n}, n_c \leq n, \\ x_c \in \mathbb{R}^{n_c}, A_c \in \mathbb{R}^{n_c \times n_c}, B_c \in \mathbb{R}^{n_c \times m}, C_c \in \mathbb{R}^{l \times n_c}, \\ R_1 \in \mathbb{R}^{n \times n}, R_1 \geq 0, R_2 \in \mathbb{R}^{m \times m}, R_2 > 0, R_{12} \in \mathbb{R}^{n \times m}. \end{aligned}$$

Furthermore, let v_1, \dots, v_p be unit-intensity, zero-mean and mutually uncorrelated white noise processes and let $w_1 \in \mathbb{R}^n$ and $w_2 \in \mathbb{R}^l$ be zero-mean white noise processes with intensities $V_1 \geq 0$ and $V_2 > 0$, respectively, and cross-intensity $V_{12} \in \mathbb{R}^{n \times l}$. Assume that v_1, w_1 and $x(0)$ are uncorrelated. We require the technical assumption that, for each i , $B_i \neq 0$ implies $C_i = 0$, i.e., the control- and measurement-dependent noises are uncorrelated.

Optimal Reduced-Order Dynamic-Compensation Problem. Given the controlled system

$$\dot{x} = (A + \sum_{i=1}^p v_i A_i)x + (B + \sum_{i=1}^p v_i B_i)u + w_1, \quad (6.1)$$

$$y = (C + \sum_{i=1}^p v_i C_i)x + w_2 \quad (6.2)$$

design an n_c th-order dynamic compensator

$$\dot{x}_c = A_c x_c + B_c y, \quad (6.3)$$

$$u = C_c x_c \quad (6.4)$$

which minimizes the performance criterion

$$J(A, B, C) = \lim_{t \rightarrow \infty} \frac{E[x^T R x + 2x^T R_{12} u + u^T R_2 u]}{t}. \quad (6.5)$$

To guarantee that J is finite and independent of initial conditions, we restrict (A_c, B_c, C_c) to the (open) set of second-moment-stabilizing triples

$$\mathcal{S} \triangleq \{(A_c, B_c, C_c) : \tilde{A}_s \otimes \tilde{A}_s + \sum_{i=1}^p \tilde{A}_i \otimes \tilde{A}_i \text{ is stable}$$

and (A_c, B_c, C_c) is controllable and observable\}

where \otimes and \odot denote Kronecker sum and product and

$$\begin{aligned} \tilde{A}_s \triangleq \begin{bmatrix} A_s & B_s C_c \\ B_c C_s & A_c \end{bmatrix}, \quad \tilde{A}_i \triangleq \begin{bmatrix} A_i & B_i C_c \\ B_c C_i & 0 \end{bmatrix}, \\ A_s \triangleq A + \frac{1}{2} \sum_{i=1}^p A_i^2, B_s \triangleq B + \frac{1}{2} \sum_{i=1}^p A_i B_i, C_s \triangleq C + \frac{1}{2} \sum_{i=1}^p C_i A_i. \end{aligned}$$

Call a square matrix positive semisimple if it has positive eigenvalues and a diagonal Jordan canonical form, i.e., if it is similar to a positive-diagonal (or, equivalently, a positive-definite) matrix. The following lemma is proved in [19].

Lemma 6.1. If $n \times n$ \hat{Q}, \hat{P} are non-negative definite and $\text{rank } \hat{Q}\hat{P} = n_c$ then there exist $n_c \times n$ G, Γ and $n_c \times n_c$ positive-semisimple M such that

$$\hat{Q}\hat{P} = G^T M \Gamma, \quad (6.6)$$

$$\Gamma G^T = I_{n_c}. \quad (6.7)$$

For convenience in stating the main result, we shall refer to G, M and Γ satisfying (6.6) and (6.7) as a projective factorization of $\hat{Q}\hat{P}$.

For convenience in stating the optimality conditions, define the following notation for

$$Q, P, \hat{Q}, \hat{P} \in \mathbb{R}^{n \times n};$$

$$R_{2s} \triangleq R_2 + \sum_{i=1}^p B_i^T (P + \hat{P}) B_i,$$

$$V_{2s} \triangleq V_2 + \sum_{i=1}^p C_i (Q + \hat{Q}) C_i^T,$$

$$Q_s \triangleq Q C_s^T + V_{12} + \sum_{i=1}^p A_i (Q + \hat{Q}) C_i^T,$$

$$P_s \triangleq B_s^T P + R_{12} + \sum_{i=1}^p B_i^T (P + \hat{P}) A_i,$$

$$A_{Qs} \triangleq A_s - Q_s V_{2s}^{-1} C_s,$$

$$A_{Ps} \triangleq A_s - B_s R_{2s}^{-1} P_s,$$

$$\tau \triangleq G^T \Gamma, \quad \tau_1 = I_n - \tau.$$

Theorem 6.1. Suppose $(A_s, B_s, C_s) \in S$ solves the optimal dynamic-compensation problem. Then there exist $n \times n$ non-negative-definite matrices Q , P , \hat{Q} and \hat{P} such that, for some projective factorization, G, M, Γ of $\hat{Q}P$, A_s , B_s , and C_s are given by

$$A_c = \Gamma(A_s - B_s R_{2s}^{-1} P_s - Q_s V_{2s}^{-1} C_s) G^T, \quad (6.8)$$

$$B_c = \Gamma Q_s V_{2s}^{-1}, \quad (6.9)$$

$$C_c = -R_{2s}^{-1} P_s G^T, \quad (6.10)$$

and such that the following conditions are satisfied:

$$0 = A_s Q + Q A_s^T + V_1 + \sum_{i=1}^p [A_i Q A_i^T + (A_i - B_i R_{2s}^{-1} P_s) \hat{Q} (A_i - B_i R_{2s}^{-1} P_s)^T] - Q_s V_{2s}^{-1} Q_s^T + \tau_1 Q_s V_{2s}^{-1} Q_s^T \tau_1, \quad (6.11)$$

$$0 = A_s^T P + P A_s + R_1 + \sum_{i=1}^p [A_i^T P A_i + (A_i - Q_s V_{2s}^{-1} C_i)^T \hat{P} (A_i - Q_s V_{2s}^{-1} C_i)] - P_s R_{2s}^{-1} P_s + \tau_1 P_s R_{2s}^{-1} P_s \tau_1, \quad (6.12)$$

$$0 = A_{Ps} \hat{Q} + \hat{Q} A_{Ps}^T + Q_s V_{2s}^{-1} Q_s^T - \tau_1 Q_s V_{2s}^{-1} Q_s^T \tau_1, \quad (6.13)$$

$$0 = A_{Qs}^T \hat{P} + \hat{P} A_{Qs} + P_s R_{2s}^{-1} P_s - \tau_1 P_s R_{2s}^{-1} P_s \tau_1, \quad (6.14)$$

$$\text{rank } \hat{Q} = \text{rank } \hat{P} = \text{rank } \hat{Q}\hat{P} = n_c.$$

Remark 6.1. Since $R_{2s} \geq R_2$, so that $R_{2s}^{-1} \leq R_2^{-1}$, it is clear that the control-dependent noise effectively suppresses the regulator gain C_c . Similarly, since $V_{2s} \geq V_2$, the measurement-dependent noise suppresses the observer gain B_c . The terms $A_i Q A_i^T$ are responsible for the decorrelation effects discussed in Section 5.

1. Lyon, R. H., Statistical Energy Analysis of Dynamical Systems: Theory and Applications, MIT Press, Cambridge, MA., 1975.
2. Hyland, D. C., "Optimal regulation of structural systems with uncertain parameters", MIT Lincoln Laboratory, TR-551, DDC# AD-A-09111/7, 1981.
3. Hyland, D. C., "Active control of large flexible spacecraft: A new design approach based on minimum information modeling of parameter uncertainties", Proc. Third VPI&SU/AIAA Symposium, pp. 631-646, Blacksburg, VA., 1981.
4. Hyland, C. C., "Optimal regulator design using minimum information modeling of parameter uncertainties: Ramifications of the new design approach", Proc. Third VPI&SU/AIAA Symposium, pp. 701-716, Blacksburg, Va., 1981.
5. Hyland, C. C., and A. N. Madiwale, "Minimum information approach to regulator design: Numerical methods and illustrative results", Proc. Third VPI&SU/AIAA Symposium, pp. 1011-118, Blacksburg, Va., 1981.
6. Hyland, D. C., and A. N. Madiwale, "A stochastic design approach for full-order compensation of structural systems with uncertain parameters", Proc. AIAA Guid. Contr. Conf., pp. 324-332, Albuquerque, NM., 1981.
7. Hyland, D. C., "Optimality conditions for fixed-order dynamic compensation of flexible spacecraft with uncertain parameters", AIAA 20th Aerospace Sciences Meeting, paper 82-0312, Orlando, FL., 1982.
8. Hyland, D. C., "Minimum information stochastic modeling of linear systems with a class of parameter uncertainties", Proc. Amer. Contr. Conf., pp. 620-627, Arlington, VA., 1982.
9. Hyland, D. C., "Maximum entropy stochastic approach to control design for uncertain structural systems", Proc. Amer. Contr. Conf., pp. 680-688, Arlington, VA., 1982.
10. Hyland, D. C., "Minimum information modeling of structural systems with uncertain parameters", Proceedings of the Workshop on Applications of Distributed System Theory to the Control of Large Space Structures, pp. 711-88, JPL, Pasadena, Ca., 1982.
11. Hyland, D. C., and A. N. Madiwale, "Fixed-order dynamic compensation through optimal projection", Proceedings of the Workshop on Applications of Distributed System Theory to the Control of Large Space Structures, pp. 409-427, JPL, Pasadena, CA., 1982.
12. Hyland, D. C., "Mean-square optimal fixed-order compensation - beyond spillover suppression", AIAA Astrodynamics Conference, paper 1403, San Diego, Ca., 1982.

13. Hyland, D. C., "The optimal projection approach to fixed-order compensation: Numerical methods and illustrative results", AIAA 21st Aerospace Sciences Meeting, paper 83-0303, Reno, NV., 1983.
14. Hyland, D. C., Mean-square optimal, full-order compensation of structural systems with uncertain parameters, MIT, Lincoln Laboratory TR-626, 1983.
15. Hyland, D. C., and D. S. Bernstein, "Explicit optimality conditions for fixed-order dynamic compensation", Proc. 22nd IEEE Conf. Dec. Contr., pp. 161-165, San Antonio, TX., 1983.
16. Hyland, D. C., "Comparison of various controller-reduction methods: Suboptimal versus optimal projection", Proc. AIAA Dynamics Specialists Conf., pp. 381-389, Palm Springs, Ca., 1984.
17. Hyland, D. C., "Application of the maximum entropy/optimal projection control design approach for large space structures", Proc. Large Space Antenna Systems Technology Conference, pp. 617-654, NASA Langley., 1984.
18. Hyland, D. C., and D. S. Bernstein, "The optimal projection equations for fixed-order dynamic compensation", IEEE Trans. Autom. Control., AC-28, pp. 1034-1037, 1982.
19. Hyland, D. C., and D. S. Bernstein, "The optimal projection approach to model reduction and the relationship between the methods of Wilson and Moore", 23rd IEEE Conf. Dec. Contr., pp. 120-126, Las Vegas, NV., 1984.
20. Bernstein, D. S., and D. C. Hyland, The optimal projection approach to designing optimal finite-dimensional controllers for distributed parameter systems, Proc. 23rd IEEE Conf. Dec. Contr., 556-560, Las Vegas, NV., 1984.
21. Davis, L. D., D. C. Hyland, and D. S. Bernstein, "Application of the maximum entropy design approach to the spacecraft control laboratory experiment (SCOLE)", Final Report, NASA Langley, 1985.
22. Bernstein, D. S., and D. C. Hyland, "The optimal projection equations for reduced-order state estimation", IEEE Trans. Autom. Contr., AC-30, pp. 583-585, 1985.
23. Jaynes, E. T., "Information theory and statistical mechanics", Phys. Rev., 106, pp. 620-630, 1957.
24. Jaynes, E. T., "New engineering applications of information theory", In J. L. Bogdanoff and F. Kozin (Eds.), proceedings of the First Symposium on Engineering Applications of Random Function Theory and Probability, Wiley, New York, pp. 163-203, 1963.
25. Jaynes, E. T., "Prior probabilities", IEEE Trans. Sys. Sci. Cybern., SSC-4, pp. 227-241, 1968.
26. Jaynes, E. T., "Where do we stand on maximum entropy?", In D. Levine and M. Tribus (Eds.), The Maximum Entropy Formalism, MIT Press, Cambridge, MA., pp. 15-118, 1979.
27. Athans, M., R. T. Ku, and S. B. Gershwin, "The uncertainty threshold principle: Some fundamental limitations of optimal decision making under dynamic uncertainty", IEEE Trans. Autom. Contr., AC-22, pp. 491-495, 1977.
28. Ku, R. J., and M. Athans, "Further results on the uncertainty threshold principle", IEEE Trans. Autom. Contr., AC-22, pp. 855-868, 1977.
29. Wong, E., and M. Zakai, "On the relation between ordinary and stochastic differential equations", Int. J. Engrg. Sci., 3, pp. 213-229, 1965.
30. Stratonovich, R. L., "A new representation for stochastic integrals", SIAM J. Contr., 4, pp. 362-371, 1966.

NUMERICAL SOLUTION OF THE OPTIMAL PROJECTION/
MAXIMUM ENTROPY DESIGN EQUATIONS FOR
LOW-ORDER, ROBUST CONTROLLER DESIGN

Dennis S. Bernstein, Lawrence D. Davis
Scott W. Greeley and David C. Hyland

Harris Corporation, GASD
MS 102/4134
Melbourne, FL 32901

ABSTRACT

This paper summarizes some recent results obtained using the optimal projection/maximum entropy control-design equations. The main results include: low-order controllers for CSDL Model #2; robust controllers for the SCOLE and VCOSS A models with modal-frequency uncertainties; and Doyle's example.

1. Introduction

The optimal projection/maximum entropy design equations are discussed in [1]-[5] and a complete, self-contained derivation appears in [5]. In brief, these equations generalize classical LQG theory in two distinct ways. First, the controller is constrained to have a fixed, reduced order and the resulting necessary conditions involve an oblique projection [1]. And, second, multiplicative white noise is introduced into the plant to capture the statistical effects of parameter uncertainty. The resulting dynamical equation is interpreted according to symmetric Stratonovich stochastic integration and, using the theory of stochastic approximation, has been motivated by the maximum entropy principle of Jaynes.

2. CSDL #2

The optimal projection (OP) reduced-order design equations were solved for the 20-state version of CSDL Model #2 treated in [6,7]. For various control-authority levels, OP designs were obtained for orders $n_c = 10, 6$ and 4 . Figure 1 summarizes the results obtained in [6,7]. Note that for compensator order $n_c = 4$, the allowable control bandwidth is severely restricted. The OP designs, however, all lie within the shaded band close to the LQG performance over a considerably expanded range of control bandwidths. Relative to LQG, the performance of the OP designs is given in Figure 2. Details of the numerical algorithm used to obtain these results are given in [8].

3. SCOLE

The SCOLE configuration is discussed at length in [9,10]. The model utilized in [9] involves 16 states, 12 actuators and 17 sensors. The LQG design reveals instability resulting from 5% modal-frequency perturbations. Using the maximum entropy (ME) design equations, a pair of controllers were obtained in the presence of stochastically modelled modal-frequency uncertainties. The first design exhibits near-LQG performance with 60% increase in robustness, while the second design is considerably more robust (behaving more like the open-loop structure) with nominal performance within 6% of LQG. These designs (Fig. 3) illustrate the performance/robustness tradeoff capabilities of the ME method. It should be noted

that, for lightly damped structures, significant modal-frequency uncertainty corresponds to pronounced spectral-resonance shifting. Frequency-domain bounds for such perturbations are consequently large and hence may result in conservative performance estimates.

4. VCOSS A

The VCOSS A model [11] is a version of CSDL model #2 involving 9 colocated sensor/actuator pairs plus 2 line-of-sight sensors. For the 28-state (14-mode) model and corresponding 28-state LQG design obtained in [11], the sensitivity to modal-frequency perturbation is shown in Figure 4. Note that instability results from 3% modal-frequency perturbations of one of the modelled poles. For the maximum entropy design (Figure 5) the robustness is considerably improved with approximately 20% performance trade. Of course, there are a continuum of intermediate designs that could be obtained for desired performance/robustness tradeoffs. The closed-loop stability margins for the full 142-mode evaluation model are shown in Figure 6.

5. Doyle's Example

As a final application of the ME design equations, we consider the problem used in [12] to demonstrate the lack of robustness of LQG designs. As shown in [12] (see [4] for notation), LQG regulators for the example

$$A = \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}, \quad B = \begin{bmatrix} 0 \\ b \end{bmatrix}, \quad C = [1 \quad 0],$$

$$V_1 = \sigma \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}, \quad V_{12} = 0, \quad V_2 = 1,$$

$$R_1 = \rho \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}, \quad R_{12} = 0, \quad R_2 = 1,$$

have arbitrarily small stability margin with regard to variations $b + \Delta b$ when σ and ρ are sufficiently large and $b = 1$. Setting $\sigma = \rho = 60$, it follows that the LQG regulator is only stable for $.93 \leq b + \Delta b \leq 1.01$. Uncertainty in b can be modelled by setting $p = 1$, $A_1 = 0$, $B_1 = [0 \quad b_1]^T$ and $C_1 = 0$. Solving the ME design equations with $b_1 = .05, .10, .15$ and $.20$ yields a series of increasingly robust controller designs with respect to both positive and negative variations Δb (see Figures 7 and 8). For more details, see [13].

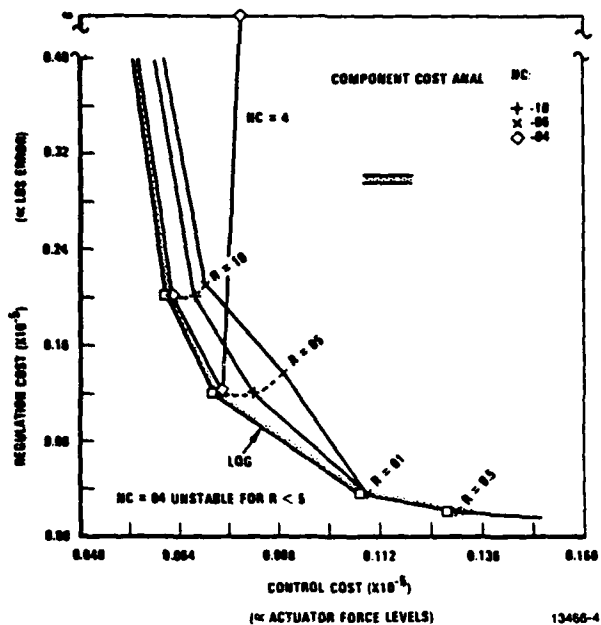


Figure 1
Optimal Projection Designs (Shaded Band)
versus Suboptimal Methods

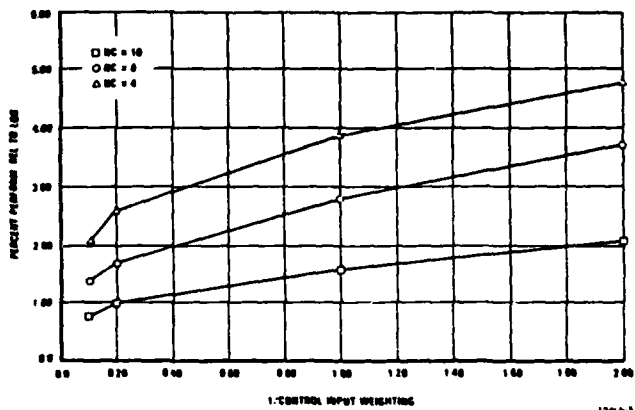


Figure 2
Optimal Projection Performance
Relative to LQG

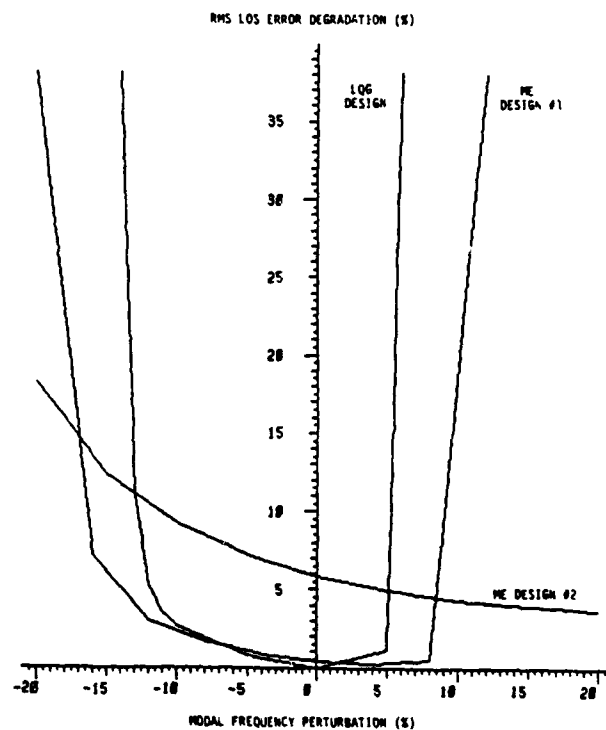


Figure 3
LQG and Maximum Entropy Designs for SCOLE:
Closed-Loop Robustness to Modal-Frequency
Perturbations

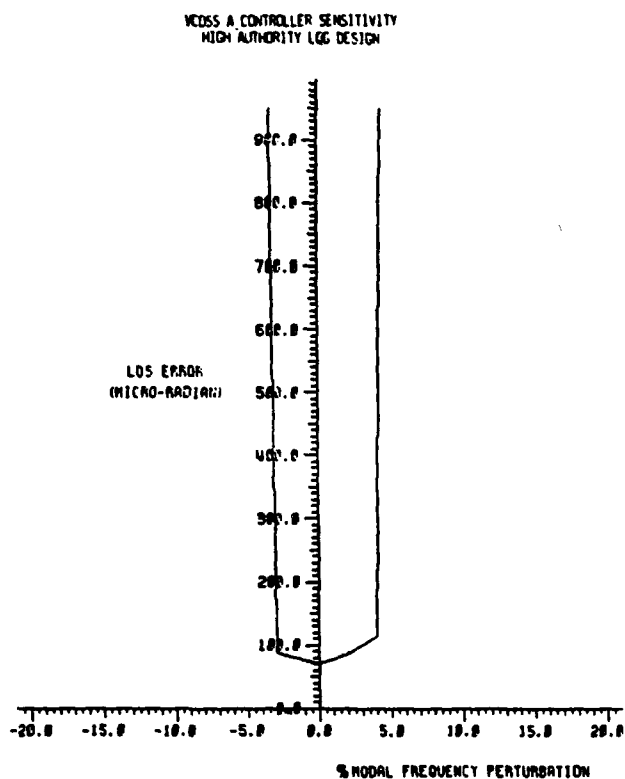


Figure 4
LQG Design for VCOSS A: Closed-Loop
Robustness to Modal-Frequency Perturbations

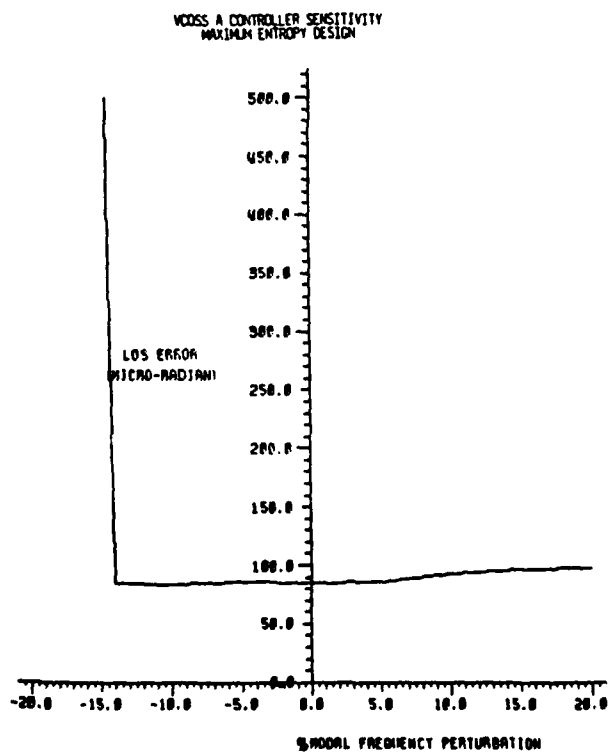


Figure 5
Maximum Entropy Design for VCOSS A:
Closed-Loop Robustness to
Modal-Frequency Perturbations

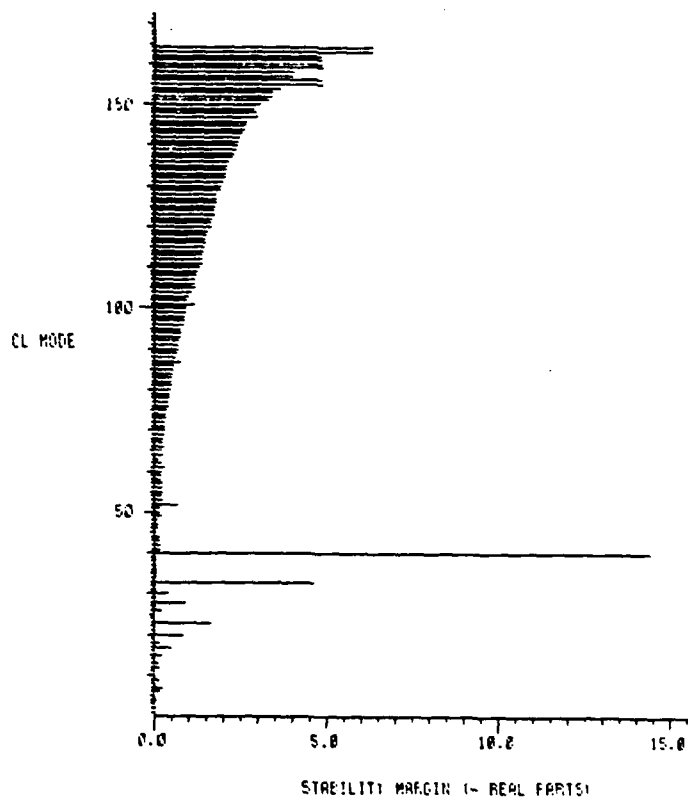


Figure 6
VCOSS A ME Design: Stability Margin
for 142-State Model

REFERENCES

1. D.C. Hyland and D. S. Bernstein, "The Optimal Projection Equations for Fixed-Order Dynamic Compensation," *IEEE Trans. Autom. Contr.*, Vol. AC-29, pp. 1034-1037, 1984.
2. D. C. Hyland, "Application of the Maximum Entropy/Optimal Projection Control Design Approach for Large Space Structures," *Proc. Large Scale Antenna Systems Technology Conference*, pp. 617-654, NASA Langley, December 1984.
3. D. S. Bernstein and D. C. Hyland, "Optimal Projection/Maximum Entropy Stochastic Modelling and Reduced-Order Design Synthesis," *Proc. IFAC Workshop on Model Error Concepts and Compensation*, Boston, MA, June 1985.
4. D. S. Bernstein and D. C. Hyland, "The Optimal Projection/Maximum Entropy Approach to Designing Low-Order, Robust Controllers for Flexible Structures," 24th IEEE Conf. Dec. Contr., Fort Lauderdale, FL, December 1985.
5. D. S. Bernstein and D. C. Hyland, "The Optimal Projection Equations for Reduced-Order Modelling, Estimation and Control of Linear Systems with Stratonovich Multiplicative White Noise," submitted for publication.
6. R. E. Skelton and P. C. Hughes, "Modal Cost Analysis for Linear Matrix Second-Order Systems," *J. Dyn. Sys. Meas. Contr.*, Vol. AC-29, pp. 520-530, 1984.
7. A. Youseff and R. E. Skelton, "Controller Reduction by Component Cost Analysis," *IEEE Trans. Autom. Contr.*, Vol. 102, pp. 151-163, 1980.
8. D. C. Hyland, "Comparison of Various Controller-Reduction Methods: Suboptimal Versus Optimal Projection," *Proc. AIAA Dynamics Specialists Conf.*, pp. 381-389, Palm Springs, CA, May 1984.
9. L. W. Taylor and A. V. Balakrishnan, "A Mathematical Problem and a Spacecraft Control Laboratory Experiment (SCOLE) Used to Evaluate Control Laws for Flexible Spacecraft ... NASA/IEEE Design Challenge," NASA/IEEE Report, June 1984.
10. L. D. Davis, D. C. Hyland and D. S. Bernstein, "Application of the Maximum Entropy Design Approach to the Spacecraft Control Laboratory Experiment (SCOLE)," Final Report, NASA Langley, January 1985.
11. J. N. Auburn, C. Z. Gregory, M. G. Lyons, R. L. Kosut and A. A. Woods, Jr., "Vibration Control of Space Structures VCOSS A: High and Low-Authority Hardware Implementation," AFWAL-TR-83-3074, Air Force Wright Aeronautical Laboratories, July 1983.
12. J. C. Doyle, "Guaranteed Margins for LQG Regulators," *IEEE Trans. Autom. Contr.*, Vol. AC-23, pp. 756-757, 1978.
13. D. S. Bernstein and S. W. Greeley, "Robust Controller Synthesis Using the Maximum Entropy Design Equations," submitted for publication.

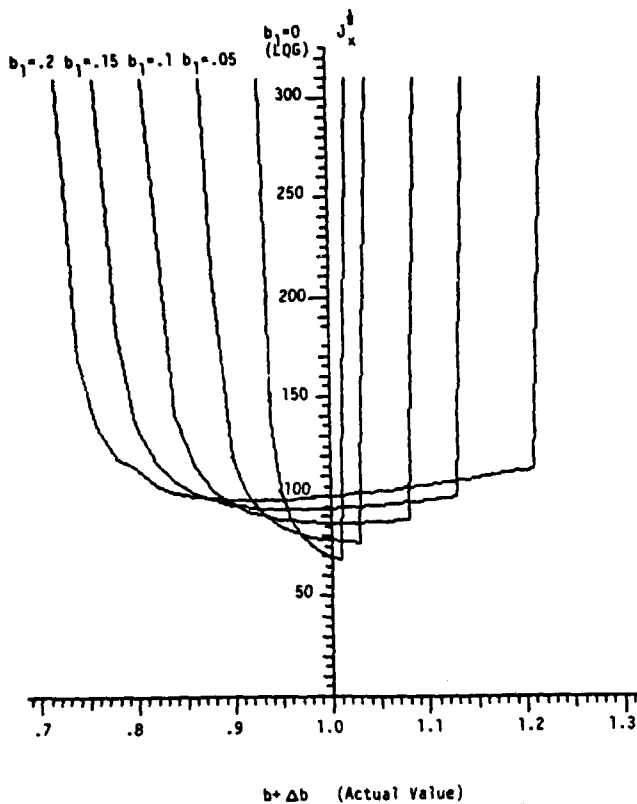


Figure 7
Robustness for Doyle's Example:
LQG plus ME Designs

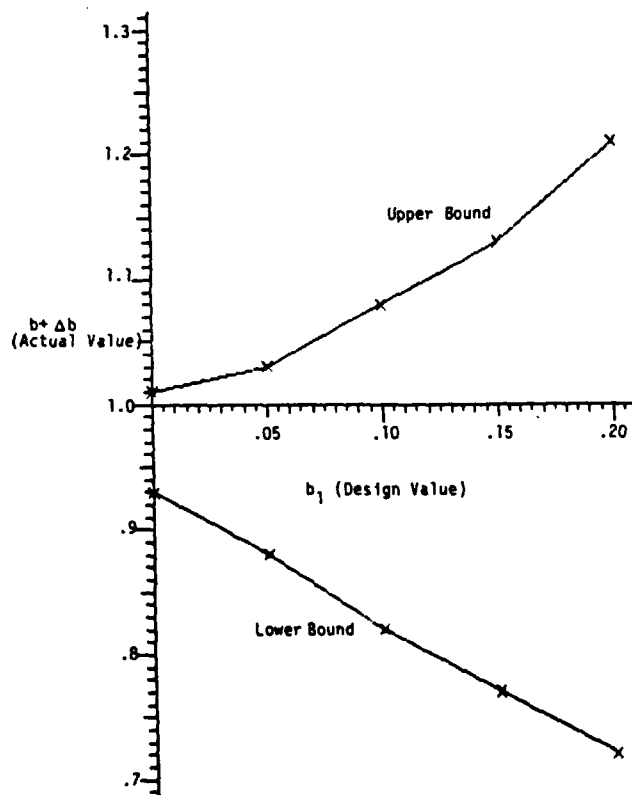


Figure 8
Stability Bounds for Doyle's Example

The Optimal Projection Equations
For Reduced-Order Modelling, Estimation and Control
Of Linear Systems With Multiplicative White Noise^{1,2}

D. S. Bernstein³ and D. C. Hyland⁴

Communicated by J. L. Speyer

¹This research was supported in part by the Air Force Office of Scientific Research under contracts F49620-84-C-0015, F49620-86-C-0002 and F49620-86-C-0038.

²We wish to thank Jill Straehla for excellent typing of the manuscript versions of the paper.

³Staff Engineer, Controls Analysis and Synthesis Group, Government Aerospace Systems Division, MS 22/4848, Harris Corporation, Melbourne, FL 32902.

⁴Group Leader, Controls Analysis and Synthesis Group, Government Aerospace Systems Division, MS 22/4848, Harris Corporation, Melbourne, Florida 32902.

Abstract. The optimal projection equations for quadratically optimal reduced-order modelling, estimation and control are generalized to include the effects of state-, control- and measurement-dependent noise.

Key Words. Feedback, control, robust, fixed-order, optimal.

1. Introduction

As is well known, LQR and LQG controllers lack guaranteed robustness with respect to arbitrary parameter variations (Refs. 1 and 2). A widely studied approach to correcting this defect involves introducing noise into the plant via the imperfectly known parameters (Refs. 3-10). Intuitively speaking, the quadratically optimal feedback controller designed in the presence of such disturbances is automatically desensitized to actual parameter variations. This was demonstrated in Ref. 11 for the example given in Ref. 1.

The contribution of the present paper is a generalization of classical steady-state LQG theory to include the effects of state-, control- and measurement-dependent noise. In contrast to the classical solution involving a pair of separated Riccati equations, the necessary conditions for quadratic optimality in the presence of multiplicative white noise consist of a system of two modified Riccati equations and two modified Lyapunov equations coupled by stochastic effects. The coupling serves as a graphic portrayal of the breakdown of the separation principle in the multiplicative noise case. When the multiplicative noise terms are set to zero, the modified Lyapunov equations drop out and the modified Riccati equations immediately reduce to the standard pair of separated LQG Riccati equations. Related results were obtained for the discrete-time, finite-interval problem in Ref. 10.

To attain further generality, a constraint is imposed on controller order as in Ref. 12. Hence, the results of the present paper also constitute a direct generalization of the coupled system of modified Riccati and Lyapunov equations which arise in characterizing reduced-order controllers.

For the special case of full-order compensation in the presence of state-dependent noise only, versions of these equations were discovered independently by Hyland (Refs. 13 and 14) and Mil'stein (Ref. 15). An interesting difference between Refs. 13 and 14 and Ref. 15 is that Mil'stein interpreted the plant model as an Ito stochastic differential equation whereas Hyland utilized the Fisk-Stratonovich definition (Refs. 16-18). In earlier work on modelling flexible mechanical structures (Refs. 19 and 20), justification for this interpretation as an appropriate model for parameter uncertainty was based

upon the Maximum Entropy Principle of Jaynes (Ref. 21) and the theory of stochastic approximation (Ref. 22). A summary of this approach and its relationship to Refs. 23 and 24 can be found in Ref. 25. Rigorous guarantees of robustness over a prescribed range of parameter variations have been obtained using Lyapunov functions (Refs. 26-29). Although the present paper utilizes an Ito model for simplicity, results based on Stratonovich models are readily obtained by means of standard transformations.

An immediate practical benefit of the structured form of the necessary conditions is the means for constructing numerical algorithms which differ fundamentally from gradient search techniques. One such iterative algorithm, proposed in Refs. 30-32, exploits the characterization of the oblique projection as the sum of rank-1 eigenprojections of the product of the rank-deficient "pseudogramians" satisfying the modified Lyapunov equations. As discussed in Ref. 32, the necessary conditions fail to specify which eigenprojections comprise the oblique projection; indeed, each choice may correspond to a local extremal. In practice, judicious choice of the eigenprojections can eliminate extremals with high cost and hence efficiently identify the global minimum. These issues are a result of the reduced-order constraint only; the stochastic effects alone do not appear to introduce extremal multiplicity.

The scope of the present paper involves deriving the optimal projection equations for reduced-order modelling, estimation and control obtained in Refs. 32, 33 and 12 to include state-, control- and measurement-dependent noise. The main results, Theorems 2.1-2.3, present the necessary conditions for optimality as systems of 2, 3, and 4 matrix equations (modified Riccati and Lyapunov equations) coupled by both the optimal projection and stochastic effects. The necessary conditions in this generality are presented here for the first time. The dynamic compensation result supports the numerical results obtained in Refs. 11 and 34. Appendix D contains the proof of Theorem 2.3; the proofs of Theorems 2.1 and 2.2 are similar and hence are omitted. Although the derivations in Refs. 32, 33 and 12 utilizing Lagrange multipliers could have been adapted to the present case, we have devised a new proof based upon Kronecker products which is thought to be more direct.

2. Problem Statement and Main Results

The following notation and definitions will be used throughout the paper.

| | |
|--|--|
| $\underline{\mathbb{R}}$ | expected value |
| $\underline{\mathbb{R}}$ | real numbers |
| $\underline{\mathbb{R}}^{\alpha \times \beta}$ | $\alpha \times \beta$ real matrices |
| $\underline{\mathbb{R}}^{\alpha}$ | $\underline{\mathbb{R}}^{\alpha \times 1}$ |
| I_{α} | $\alpha \times \alpha$ identity matrix |
| $z_{(i)}$ | i th element of vector z |
| $z_{(i,j)}$ | (i,j) element of matrix z |
| z^T | transpose of vector or matrix z |
| z^{-T} | $(z^T)^{-1}$ or $(z^{-1})^T$ |
| $\rho(z)$ | rank of matrix z |
| $\text{tr } z$ | trace of square matrix z |
| $\ z\ $ | $(\text{tr } zz^T)^{\frac{1}{2}}$ (Frobenius norm) |
| $\text{diag}(a_1, \dots, a_{\alpha})$ | $\alpha \times \alpha$ diagonal matrix with listed diagonal elements |
| E_i | matrix with unity in the (i,i) position and zeros elsewhere |
| $\Pi_i(\Psi)$ | $\Psi E_i \Psi^{-1}$ |
| $x \otimes y$ | $\begin{bmatrix} x_{(1,1)}y \dots x_{(1,\beta)}y \\ \vdots \\ x_{(\alpha,1)}y \dots x_{(\alpha,\beta)}y \end{bmatrix}$ $x \in \underline{\mathbb{R}}^{\alpha \times \beta}, y \in \underline{\mathbb{R}}^{\gamma \times \delta}$ (Kronecker product. Refs. 35 and 36) |

$X \oplus Y$

$X \otimes I_\beta + I_\alpha \otimes Y$
 $X \in \mathbb{R}^{\alpha \times \alpha}, Y \in \mathbb{R}^{\beta \times \beta}$ (Kronecker sum)

Z^\sharp

group generalized inverse (Ref. 37)

$\text{row}_i(Z)$

i th row of matrix Z

$\text{col}_i(Z)$

i th column of matrix Z

$\text{vec}(Z)$

$\begin{bmatrix} \text{col}_1(Z) \\ \vdots \\ \text{col}_\beta(Z) \end{bmatrix} \in \mathbb{R}^{\alpha\beta}, Z \in \mathbb{R}^{\alpha \times \beta}$

$\text{vec}_{(\alpha,\beta)}^{-1}(Z)$

$\begin{bmatrix} z_{(1)} & \dots & z_{(\alpha\beta-\alpha+1)} \\ \vdots & & \vdots \\ z_{(\alpha)} & \dots & z_{(\alpha\beta)} \end{bmatrix} \in \mathbb{R}^{\alpha \times \beta}, Z \in \mathbb{R}^{\alpha \times \beta}$

stable matrix

matrix with eigenvalues in open left half plane

nonnegative-definite matrix

symmetric matrix with nonnegative eigenvalues ($Z \geq 0$)

positive-definite matrix

symmetric matrix with positive eigenvalues ($Z > 0$)

semisimple eigenvalue

eigenvalue with equal algebraic and geometric multiplicity

simple eigenvalue

eigenvalue with unity algebraic multiplicity

group-invertible matrix

matrix Z satisfying $\rho(Z) = \rho(Z^2)$, i.e., matrix which is either invertible or whose zero eigenvalue is semisimple (Ref. 37)

semisimple matrix

matrix with semisimple eigenvalues (i.e., nondefective matrix)

real-semisimple matrix

semisimple matrix with real eigenvalues

nonnegative-semisimple matrix

semisimple matrix with nonnegative eigenvalues

positive-semisimple matrix

semisimple matrix with positive eigenvalues

simple matrix

matrix with distinct (i.e., simple) eigenvalues

r

generic subscript denoting m , e or c

$n, m, l, n_m, n_e, n_c, k, p, q$

positive integers; $1 \leq n_m, n_e, n_c \leq n$; $l, m \leq q$

\tilde{n}_r

$n + n_r$

x, x_m, x_e, x_c

n, n_m, n_e, n_c -dimensional vectors

u, y, y_e

m, l, k -dimensional vectors

A, A_1, \dots, A_p

$n \times n$ matrices

B, B_1, \dots, B_p

$n \times m$ matrices

C, C_1, \dots, C_p

$l \times n$ matrices

A_m, B_m, C_m

$n_m \times n_m, n_m \times m, l \times n_m$ matrices

A_e, B_e, C_e

$n_e \times n_e, n_e \times l, k \times n_e$ matrices

A_c, B_c, C_c

$n_c \times n_c, n_c \times l, m \times n_c$ matrices

R, N, R_2

$l \times l, k \times k, m \times m$ positive-definite matrices

R_1

$n \times n$ nonnegative-definite matrix

R_{12}

$n \times m$ matrix; $R_1 - R_{12} R_2^{-1} R_{12}^T \geq 0$

\tilde{R}

$$\begin{bmatrix} R_1 & R_{12} \\ R_{12}^T & R_2 \end{bmatrix} \geq 0$$

L

$k \times n$ matrix

$v_{1t}, \dots, v_{pt}, w_{1t}, \dots, w_{qt}$

standard, independent Wiener processes, $t \geq 0$

w_t

$(w_{1t}, \dots, w_{qt})^T$

G, G_1, G_2

$m \times q, n \times q, l \times q$ matrices; $\rho(G) = m, \rho(G_2) = l$

$$v_1 \triangleq G_1 G_1^T \geq 0, \quad v_{12} \triangleq G_1 G_2^T, \quad v_2 \triangleq G_2 G_2^T > 0, \quad v \triangleq G G^T > 0,$$

$$\tilde{G} \triangleq \begin{bmatrix} G_1 \\ G_2 \end{bmatrix}, \quad \tilde{V} \triangleq \tilde{G} \tilde{G}^T = \begin{bmatrix} v_1 & v_{12} \\ v_{12}^T & v_2 \end{bmatrix} \geq 0,$$

$$\tilde{B}_m \triangleq \begin{bmatrix} B \\ B_m \end{bmatrix}, \quad \tilde{B}_e \triangleq \begin{bmatrix} I_n & 0 \\ 0 & B_e \end{bmatrix}, \quad \tilde{B}_c \triangleq \begin{bmatrix} I_n & 0 \\ 0 & B_c \end{bmatrix},$$

$$\tilde{G}_m \triangleq \tilde{B}_m G, \quad \tilde{G}_e \triangleq \tilde{B}_e G, \quad \tilde{G}_c \triangleq \tilde{B}_c G,$$

$$\tilde{V}_r \triangleq \tilde{G}_r \tilde{G}_r^T, \quad \tilde{V}_m \triangleq \tilde{B}_m \tilde{V} \tilde{B}_m^T, \quad \tilde{V}_e \triangleq \tilde{B}_e \tilde{V} \tilde{B}_e^T, \quad \tilde{V}_c \triangleq \tilde{B}_c \tilde{V} \tilde{B}_c^T,$$

$$\tilde{C}_m \triangleq [C \quad -C_m], \quad \tilde{C}_e \triangleq [L \quad -C_e], \quad \tilde{C}_c \triangleq \begin{bmatrix} I_n & 0 \\ 0 & C_c \end{bmatrix},$$

$$\tilde{R}_m \triangleq \tilde{C}_m^T R \tilde{C}_m, \quad \tilde{R}_e \triangleq \tilde{C}_e^T N \tilde{C}_e, \quad \tilde{R}_c \triangleq \tilde{C}_c^T R \tilde{C}_c,$$

$$\underline{A} \triangleq A \oplus A + \sum_{i=1}^P A_i \otimes A_i, \quad \tilde{x}_r = \begin{bmatrix} x \\ x_r \end{bmatrix},$$

$$\begin{aligned} \tilde{A}_m &\triangleq \begin{bmatrix} A & 0 \\ 0 & A_m \end{bmatrix}, & \tilde{A}_e &\triangleq \begin{bmatrix} A & 0 \\ B_e C & A_e \end{bmatrix}, & \tilde{A}_c &\triangleq \begin{bmatrix} A & B C_c \\ B_c C & A_c \end{bmatrix}, \\ \tilde{A}_{mi} &\triangleq \begin{bmatrix} A_i & 0 \\ 0 & 0 \end{bmatrix}, & \tilde{A}_{ei} &\triangleq \begin{bmatrix} A_i & 0 \\ B_e C_i & 0 \end{bmatrix}, & \tilde{A}_{ci} &\triangleq \begin{bmatrix} A_i & B_i C_c \\ B_c C_i & 0 \end{bmatrix}, \end{aligned}$$

$$\tilde{A}_r = \tilde{A}_r \oplus \tilde{A}_r + \sum_{i=1}^P \tilde{A}_{ri} \otimes \tilde{A}_{ri}.$$

For the following definitions, let $Q, P, \hat{Q}, \hat{P} \in \mathbb{R}^{n \times n}$:

$$\hat{R}_2 \triangleq R_2 + \sum_{i=1}^P B_i^T (P + \hat{P}) B_i,$$

$$\hat{V}_2 \triangleq V_2 + \sum_{i=1}^P C_i (Q + \hat{Q}) C_i^T,$$

$$Q \triangleq Q C^T + V_{12} + \sum_{i=1}^P A_i (Q + \hat{Q}) C_i^T,$$

$$P \triangleq B^T P + R_{12}^T + \sum_{i=1}^P B_i^T (P + \hat{P}) A_i,$$

$$A_Q \triangleq A - Q \hat{V}_2^{-1} C,$$

$$A_P \triangleq A - B \hat{R}_2^{-1} P.$$

Using the above notation we can state the reduced-order modelling, estimation and control problems.

Reduced-Order Modelling Problem. Given the n th-order model

$$dx_t = A x_t dt + \sum_{i=1}^P A_i x_t dv_{it} + G_1 dw_t, \quad (1)$$

$$y_t = C x_t, \quad (2)$$

where $t \in [0, \infty)$, determine an n_m th-order model

$$dx_{mt} = A_m x_{mt} dt + B_m G dw_t, \quad (3)$$

$$y_{mt} = C_m x_{mt}, \quad (4)$$

which minimizes the model-reduction criterion

$$J_m(A_m, B_m, C_m) \triangleq \limsup_{t \rightarrow \infty} \mathbb{E}[(y_t - y_{mt})^T R (y_t - y_{mt})]. \quad (5)$$

Reduced-Order State-Estimation Problem. Given the nth-order observed system

$$dx_t = Ax_t dt + \sum_{i=1}^P A_i x_t dv_{it} + BGdw_t, \quad (6)$$

$$dy_t = Cx_t dt + \sum_{i=1}^P C_i x_t dv_{it} + G_2 dw_t, \quad (7)$$

where $t \in [0, \infty)$, design an n_e th-order state estimator

$$dx_{et} = A_e x_{et} dt + B_e dy_t, \quad (8)$$

$$y_{et} = C_e x_{et}, \quad (9)$$

which minimizes the state-estimation criterion

$$J_e(A_e, B_e, C_e) \triangleq \limsup_{t \rightarrow \infty} \mathbb{E}[(Lx_t - y_{et})^T N(Lx_t - y_{et})]. \quad (10)$$

Reduced-Order Dynamic-Compensation Problem. Given the nth-order observed and controlled system

$$dx_t = Ax_t dt + \sum_{i=1}^P A_i x_t dv_{it} + Bu_t dt + \sum_{i=1}^P B_i u_t dv_{it} + G_1 dw_t, \quad (11)$$

$$dy_t = Cx_t dt + \sum_{i=1}^P C_i x_t dv_{it} + G_2 dw_t, \quad (12)$$

where $t \in [0, \infty)$, design an n_c th-order dynamic compensator

$$dx_{ct} = A_c x_{ct} dt + B_c dy_t, \quad (13)$$

$$u_t = C_c x_{ct}, \quad (14)$$

which minimizes the dynamic-compensation criterion

$$J_c(A_c, B_c, C_c) \triangleq \limsup_{t \rightarrow \infty} \mathbb{E}[x_t^T R_1 x_t + 2x_t^T R_{12} u_t + u_t^T R_2 u_t]. \quad (15)$$

Clearly, J_m , J_e and J_c are nonnegative, extended-real-valued functionals defined on appropriate Euclidean spaces. Explicit expressions for these functionals are now given. Henceforth we assume that $\mathbb{E} \|\bar{x}_{r0}\|^2 < \infty$ and that \bar{x}_{r0} and $v_{1t}, \dots, v_{pt}, w_t$ are uncorrelated, $t \geq 0$.

Proposition 2.1. The nonnegative-definite covariance $\bar{Q}_r(t) \triangleq \mathbb{E}[\bar{x}_{rt} \bar{x}_{rt}^T]$, $t \geq 0$, is given by

$$\dot{\bar{Q}}_r(t) = \bar{A}_r \bar{Q}_r(t) + \bar{Q}_r(t) \bar{A}_r^T + \sum_{i=1}^p \bar{A}_{ri} \bar{Q}_r(t) \bar{A}_{ri}^T + \bar{V}_r, \quad t \geq 0, \quad (16)$$

or, explicitly, by

$$\bar{Q}_r(t) = \text{vec}_{[n_r, n_r]}^{-1} \left(e^{\bar{A}_r t} \text{vec } \bar{Q}_r(0) + \int_0^t e^{\bar{A}_r \sigma} d\sigma \text{vec } \bar{V}_r \right). \quad (17)$$

The cost criteria J_m , J_e and J_c are given by

$$J_r(A_r, B_r, C_r) = \limsup_{t \rightarrow \infty} \text{tr } \bar{Q}_r(t) \bar{R}_r. \quad (18)$$

or, equivalently, by

$$J_r(A_r, B_r, C_r) = \limsup_{t \rightarrow \infty} (\text{vec } \bar{R}_r)^T \left(e^{\bar{A}_r t} \text{vec } \bar{Q}_r(0) + \int_0^t e^{\bar{A}_r \sigma} d\sigma \text{vec } \bar{V}_r \right). \quad (19)$$

Proof. See Appendix A. □

The finiteness and smoothness of J_m , J_e and J_c clearly depend upon the interrelationships among $\tilde{Q}_r(0)$, \tilde{A}_r , \tilde{R}_r and \tilde{V}_r . To avoid a detailed analysis and to guarantee that J_m , J_e , and J_c are finite and independent of initial data, we restrict our consideration to second-moment stable or second-moment stabilizing design triples. Furthermore, to avoid degeneracy in later developments (and without loss of generality) only minimal (i.e., controllable and observable) realizations are admitted. Hence, for the modelling, estimation and control problems define the open sets

$$\underline{S}_r = \{(A_r, B_r, C_r) : \tilde{A}_r \text{ is stable and } (A_r, B_r, C_r) \text{ is minimal}\}.$$

In the following result we abuse notation slightly and let \tilde{Q}_r denote $\lim_{t \rightarrow \infty} \tilde{Q}_r(t)$.

Proposition 2.2. Suppose \underline{S}_r is nonempty. If $(A_r, B_r, C_r) \in \underline{S}_r$ then $\tilde{Q}_r \stackrel{\Delta}{=} \lim_{t \rightarrow \infty} \tilde{Q}_r(t)$ exists and is given by the unique, nonnegative-definite solution to

$$0 = \tilde{A}_r \tilde{Q}_r + \tilde{Q}_r \tilde{A}_r^T + \sum_{i=1}^p \tilde{A}_{ri} \tilde{Q}_r \tilde{A}_{ri}^T + \tilde{V}_r. \quad (20)$$

or, explicitly, by

$$\tilde{Q}_r = \text{vec}_{(n_r, n_r)}^{-1} \begin{pmatrix} -\tilde{A}_r^{-1} \text{vec } \tilde{V}_r \end{pmatrix}. \quad (21)$$

Hence,

$$J_r(A_r, B_r, C_r) = \text{tr } \tilde{Q}_r \tilde{R}_r. \quad (22)$$

or, equivalently,

$$J_r(A_r, B_r, C_r) = -(\text{vec } \tilde{R}_r)^T \tilde{A}_r^{-1} \text{vec } \tilde{V}_r. \quad (23)$$

Proof. See Appendix A. □

As a side note we examine the evolution of the mean value of \tilde{x}_r .

Proposition 2.3. The mean

$$\tilde{m}(t) \triangleq \mathbb{E} \tilde{x}_{rt}, \quad t \geq 0, \quad (24)$$

satisfies

$$\dot{\tilde{m}}(t) = \tilde{A}_r \tilde{m}(t), \quad t \geq 0. \quad (25)$$

Furthermore, if $(A_r, B_r, C_r) \in \underline{S}_r$ then \tilde{A}_r is stable and thus

$$\lim_{t \rightarrow \infty} \tilde{m}(t) = 0. \quad (26)$$

Proof. See Appendix A. □

Of course, it is useful to know when the sets \underline{S}_m , \underline{S}_e and \underline{S}_c are nonempty. Although for the closed-loop control problem the question is complex because of stabilizability concerns, the modelling and estimation problems permit considerable simplification.

Proposition 2.4. \underline{S}_m (alternatively, \underline{S}_e) is nonempty if and only if \underline{A} is stable. In this case \underline{S}_m and \underline{S}_e are given by

$$\underline{S}_m = \{(A_m, B_m, C_m): A_m \text{ is stable and } (A_m, B_m, C_m) \text{ is minimal}\},$$

$$\underline{S}_e = \{(A_e, B_e, C_e): A_e \text{ is stable and } (A_e, B_e, C_e) \text{ is minimal}\}.$$

Proof. See Appendix B. □

The following observation concerns the smoothness of the cost functionals.

Proposition 2.5. The functionals J_r are infinitely Frechet differentiable on \underline{S}_r .

Proof. From Lemma 3.7.2, p. 203 of Ref. 38, it follows that the map $W \rightarrow W^{-1}$ defined on the set of invertible matrices is C^∞ . The result follows from the chain rule and (23). □

It is now possible to proceed with the principal aim of the paper which is to characterize solutions of the reduced-order modelling, estimation and control problems by means of a first-order variational analysis. To this end, one additional assumption is required. In order to obtain closed-form expressions for extremal values of the closed-loop control gains, the dynamic-compensation problem requires the technical assumption

$$[B_i \neq 0 \Rightarrow C_i = 0], i=1, \dots, p, \quad (27)$$

or, equivalently,

$$[C_i \neq 0 \Rightarrow B_i = 0], i=1, \dots, p, \quad (28)$$

i.e., for each $i \in \{1, \dots, p\}$, B_i and C_i are not both nonzero. Essentially, (27) expresses the condition that the control-dependent and measurement-dependent noises are independent. There are no constraints, however, on correlations with the state-dependent noise.

In order to state the main results we require some additional notation and a lemma concerning a pair of nonnegative-definite matrices. For a real-semisimple matrix $X \in \mathbb{R}^{n \times n}$ define the set of diagonalizing matrices

$$\underline{D}(X) \triangleq \{\Psi \in \mathbb{R}^{n \times n}: \Psi^{-1}X\Psi \text{ is diagonal}\},$$

and, for a pair of nonnegative-definite matrices $X, Y \in \mathbb{R}^{n \times n}$ define the set of contragrediently diagonalizing matrices

$$\underline{C}(X, Y) \triangleq \{\Psi \in \mathbb{R}^{n \times n}: \Psi^{-1}X\Psi^{-T} \text{ and } \Psi^TY\Psi \text{ are diagonal}\}$$

and the subset of balancing transformations

$$\underline{B}(X, Y) \triangleq \{\Psi \in \underline{C}(X, Y): \Psi^{-1}X\Psi^{-T} = \Psi^TY\Psi\}.$$

The following result unifies and extends similar results found in Refs. 32, 33 and 12.

Lemma 2.1. Suppose $\hat{Q}, \hat{P} \in \mathbb{R}^{n \times n}$ are nonnegative definite and $\rho(\hat{Q}\hat{P}) = n_r$. Then the following statements hold:

- (i) $\emptyset \neq \underline{C}(\hat{Q}, \hat{P}) \subset \underline{D}(\hat{Q}\hat{P})$.
- (ii) $\hat{Q}\hat{P}$ is nonnegative semisimple.
- (iii) The $n \times n$ matrix

$$\tau \triangleq \hat{Q}\hat{P}(\hat{Q}\hat{P})^\# \quad (29)$$

is idempotent, i.e., τ is an oblique projection.

(iv) There exists $\Psi \in \underline{C}(\hat{Q}, \hat{P})$ with $(\Psi^{-1}\hat{Q}\hat{P}\Psi)_{(i,i)} \neq 0$, $i=1, \dots, n_r$, such that τ is given by

$$\tau = \sum_{i=1}^{n_r} \Pi_i(\Psi). \quad (30)$$

- (v) If $\rho(\hat{Q}) = \rho(\hat{P}) = n_r$ then $\underline{B}(\hat{Q}, \hat{P}) \neq \emptyset$.

(vi) If $\rho(\hat{Q}) = \rho(\hat{P}) = n_r$ then there exists $\Psi \in B(\hat{Q}, \hat{P})$ with $(\Psi^{-1} \hat{Q} \hat{P} \Psi)_{(i,i)} \neq 0, i=1, \dots, n_r$, such that τ is given by (30).

(vii) If $\rho(\hat{Q}) = \rho(\hat{P}) = n_r$ then

$$\hat{Q} = \tau \hat{Q} = \hat{Q} \tau^T = \tau \hat{Q} \tau^T, \quad (31)$$

$$\hat{P} = \tau^T \hat{P} = \hat{P} \tau = \tau^T \hat{P} \tau. \quad (32)$$

(viii) There exist $G, \Gamma \in \underline{\mathbb{R}}^{n_r \times n_r}$ and positive-semisimple $M \in \underline{\mathbb{R}}^{n_r \times n_r}$ such that

$$\hat{\hat{Q}} \hat{\hat{P}} = G^T M \Gamma, \quad (33)$$

$$\Gamma G^T = I_{n_r}. \quad (34)$$

(ix) If $\bar{G}, \bar{\Gamma} \in \underline{\mathbb{R}}^{n_r \times n_r}$ and $\bar{M} \in \underline{\mathbb{R}}^{n_r \times n_r}$ satisfy

$$\hat{\hat{Q}} \hat{\hat{P}} = \bar{G}^T \bar{M} \bar{\Gamma}, \quad (35)$$

$$\bar{\Gamma} \bar{G}^T = I_{n_r}, \quad (36)$$

then there exists invertible $S \in \underline{\mathbb{R}}^{n_r \times n_r}$ such that $\bar{G} = S^{-T} G$, $\bar{\Gamma} = S \Gamma$ and $\bar{M} = S M S^{-1}$.

(x) If $G, \Gamma \in \underline{\mathbb{R}}^{n_r \times n_r}$ and $M \in \underline{\mathbb{R}}^{n_r \times n_r}$ satisfy (33) and (34) then M is invertible, $(\hat{\hat{Q}} \hat{\hat{P}})^{\#} = G^T M^{-1} \Gamma$ and

$$\tau = G^T \Gamma. \quad (37)$$

Proof. See Appendix C. □

For convenience we shall call $G, \Gamma \in \mathbb{R}^{n \times n}$ and $M \in \mathbb{R}^{n \times n}$ satisfying (33) and (34) a projective factorization of $\hat{\hat{Q}}\hat{\hat{P}}$. Furthermore, define the complementary oblique projection

$$\tau_1 \triangleq I_n - \tau \quad (38)$$

and let $J'_r(A_r, B_r, C_r)$ denote the Frechet derivative of J_r evaluated at (A_r, B_r, C_r) .

It is now possible to state the main results, which provide a parameterization of triples $(A_r, B_r, C_r) \in \underline{S}_r$ for which the first Frechet derivative of J_r vanishes.

Theorem 2.1. Assume \underline{A} is stable. Then, for $(A_m, B_m, C_m) \in \underline{S}_m$,

$$J'_m(A_m, B_m, C_m) = 0 \quad (39)$$

if and only if there exist $n \times n$ nonnegative-definite matrices \hat{Q} and \hat{P} such that, for some projective factorization G, M, Γ of $\hat{\hat{Q}}\hat{\hat{P}}$, A_m, B_m and C_m are given by

$$A_m = \Gamma A G^T, \quad (40)$$

$$B_m = \Gamma B, \quad (41)$$

$$C_m = C G^T, \quad (42)$$

and such that, with $\tau \triangleq \hat{\hat{Q}}\hat{\hat{P}}(\hat{\hat{Q}}\hat{\hat{P}})^\# = G^T \Gamma$ and $\tau_1 \triangleq I_n - \tau$, the following conditions are satisfied:

$$0 = \hat{A}\hat{Q} + \hat{Q}\hat{A}^T + BVB^T - \tau_1 BVB^T \tau_1^T. \quad (43)$$

$$0 = \hat{A}^T \hat{P} + \hat{P}\hat{A} + C^T RC - \tau_1^T C^T RC \tau_1. \quad (44)$$

$$\rho(\hat{Q}) = \rho(\hat{P}) = \rho(\hat{Q}\hat{P}) = n_m. \quad (45)$$

Furthermore, if $(A_m, B_m, C_m) \in \underline{S}_m$ satisfies (39) then the extremal cost is given by

$$\begin{aligned} J_m(A_m, B_m, C_m) &= \text{tr}[(W_c - \hat{Q})C^T RC] \\ &= \text{tr}[(W_o - \hat{P})BVB^T] \\ &= 2\text{tr}[(\hat{Q}\hat{P} - W_c W_o)A] - 2 \sum_{i=1}^P \text{tr} W_c A_i^T W_o A_i. \end{aligned} \quad (46)$$

where $W_c, W_o \in \mathbb{R}^{n \times n}$ are the unique, nonnegative-definite solutions to

$$0 = AW_c + W_c A^T + \sum_{i=1}^P A_i W_c A_i^T + BVB^T, \quad (47)$$

$$0 = A^T W_o + W_o A + \sum_{i=1}^P A_i^T W_o A_i + C^T RC. \quad (48)$$

Theorem 2.2. Assume \underline{A} is stable. Then, for $(A_e, B_e, C_e) \in \underline{S}_e$.

$$J_e'(A_e, B_e, C_e) = 0 \quad (49)$$

if and only if there exist $n \times n$ nonnegative-definite matrices Q, \hat{Q} and \hat{P} such that, for some projective factorization G, M, Γ of $\hat{Q}\hat{P}$, A_e, B_e and C_e are given by

$$A_e = \Gamma(A - Q\hat{V}_2^{-1}C)G^T. \quad (50)$$

$$B_e = \Gamma Q\hat{V}_2^{-1}. \quad (51)$$

$$C_e = LG^T, \quad (52)$$

and such that, with $\tau \triangleq \hat{Q}\hat{P}(\hat{Q}\hat{P})^\# = G^T\Gamma$ and $\tau_1 \triangleq I_n - \tau$, the following conditions are satisfied:

$$0 = AQ + QA^T + V_1 + \sum_{i=1}^p A_i(Q + \hat{Q})A_i^T - Q\hat{V}_2^{-1}Q^T + \tau_1 Q\hat{V}_2^{-1}Q^T\tau_1^T. \quad (53)$$

$$0 = A\hat{Q} + \hat{Q}A^T + Q\hat{V}_2^{-1}Q^T - \tau_1 Q\hat{V}_2^{-1}Q^T\tau_1^T. \quad (54)$$

$$0 = A_Q^T\hat{P} + \hat{P}A_Q + L^T N L - \tau_1^T L^T N L \tau_1. \quad (55)$$

$$\rho(\hat{Q}) = \rho(\hat{P}) = \rho(\hat{Q}\hat{P}) = n_e. \quad (56)$$

Furthermore, if $(A_e, B_e, C_e) \in \underline{S}_e$ satisfies (49) then the extremal cost is given by

$$J_e(A_e, B_e, C_e) = \text{tr } QL^T N L. \quad (57)$$

Theorem 2.3. Assume (27) holds and \underline{S}_c is nonempty. Then for $(A_c, B_c, C_c) \in \underline{S}_c$,

$$J'_c(A_c, B_c, C_c) = 0 \quad (58)$$

if and only if there exist $n \times n$ nonnegative-definite matrices Q, P, \hat{Q} and \hat{P} such that, for some projective factorization G, M, Γ of $\hat{Q}\hat{P}$, A_c, B_c and C_c are given by

$$A_c = \Gamma(A - B\hat{R}_2^{-1}\hat{P} - Q\hat{V}_2^{-1}C)G^T, \quad (59)$$

$$B_c = \Gamma Q\hat{V}_2^{-1}, \quad (60)$$

$$C_c = -\hat{R}_2^{-1}\hat{P}G^T, \quad (61)$$

and such that, with $r \triangleq \hat{Q}\hat{P}(\hat{Q}\hat{P})^\# = G^T\Gamma$ and $\tau_1 \triangleq I_n - \tau$, the following conditions are satisfied:

$$0 = A\hat{Q} + \hat{Q}A^T + V_1 + \sum_{i=1}^P [A_i \hat{Q} A_i^T + (A_i - B_i \hat{R}_2^{-1} \hat{P}) \hat{Q} (A_i - B_i \hat{R}_2^{-1} \hat{P})^T] - Q\hat{V}_2^{-1}Q^T + \tau_1 Q\hat{V}_2^{-1}Q^T \tau_1^T. \quad (62)$$

$$0 = A^T \hat{P} + \hat{P}A + R_1 + \sum_{i=1}^P [A_i^T \hat{P} A_i + (A_i - Q\hat{V}_2^{-1}C_i)^T \hat{P} (A_i - Q\hat{V}_2^{-1}C_i)] - \hat{P}^T \hat{R}_2^{-1} \hat{P} + \tau_1^T \hat{P}^T \hat{R}_2^{-1} \hat{P} \tau_1. \quad (63)$$

$$0 = A_P \hat{Q} + \hat{Q} A_P^T + Q\hat{V}_2^{-1}Q^T - \tau_1 Q\hat{V}_2^{-1}Q^T \tau_1^T. \quad (64)$$

$$0 = A_Q^T \hat{P} + \hat{P} A_Q + \hat{P}^T \hat{R}_2^{-1} \hat{P} - \tau_1^T \hat{P}^T \hat{R}_2^{-1} \hat{P} \tau_1. \quad (65)$$

$$\rho(\hat{Q}) = \rho(\hat{P}) = \rho(\hat{Q}\hat{P}) = n_c. \quad (66)$$

Furthermore, if $(A_c, B_c, C_c) \in \underline{S}_c$ satisfies (58) then the extremal cost is given by

$$\begin{aligned} \hat{J}_c(A_c, B_c, C_c) = & \text{tr}[(Q + \hat{Q})R_1 - 2R_{12}\hat{R}_2^{-1}\hat{P}\hat{Q} \\ & + \hat{P}^T \hat{R}_2^{-1} R_2 \hat{R}_2^{-1} \hat{P} \hat{Q}] \end{aligned} \quad (67)$$

3. Appendix A: Proof of Propositions 2.1, 2.2 and 2.3

First note that (1)-(4), (6)-(9) and (11)-(14) can be written as

$$d\tilde{x}_{rt} = \tilde{A}_r \tilde{x}_{rt} dt + \sum_{i=1}^P \tilde{A}_{ri} \tilde{x}_{rt} dv_{it} + \tilde{G}_r dw_t. \quad (68)$$

From Theorem 8.5.5, p. 142 of Ref. 17 (or from the Ito differential rule) it follows that the nonnegative-definite covariance $\tilde{Q}_r(t)$ is given by (16). Furthermore, (5), (10) and (15) are equivalent to (18). Rewriting $\tilde{Q}_r(t)$ in the form (see Refs. 35 and 36)

$$\dot{\tilde{Q}}_r(t) = \tilde{A}_r \tilde{Q}_r(t) + \text{vec } \tilde{V}_r \quad (69)$$

leads to (19). □

To prove Proposition 2.2, note that the stability of \tilde{A}_r implies, by (69), that $\tilde{Q}_r \triangleq \lim_{t \rightarrow \infty} \tilde{Q}_r(t)$ exists and is given by (21), which satisfies (20). Clearly, $\tilde{Q}_r \geq 0$ since $\tilde{Q}_r(t) \geq 0$, $t \geq 0$. Now (22) and (23) follow from (18) and (19). □

To prove Proposition 2.3, note that the differential equation for $\tilde{m}(t)$ is an immediate consequence of (68). To show that \tilde{A}_r is stable, we proceed as in Lemma 2.2 of Ref. 4. Repeating the steps leading to (22) with \tilde{V}_r replaced by $I_{n_r}^-$, it follows (see Ref. 4) that $(\tilde{A}_r, I_{n_r}^-)$ is stabilizable. Hence,

$$\int_0^t e^{\tilde{A}_r \sigma} \tilde{A}_r^T e^{\tilde{A}_r^T \sigma} d\sigma$$

is bounded as $t \rightarrow \infty$, which implies \tilde{A}_r is stable. □

4. Appendix B: Proof of Proposition 2.4

We require some elementary properties of the Kronecker product (Refs. 35 and 36) applied to partitioned matrices. For $X \in \underline{\mathbb{R}}^{s \times t}$ and $Y \in \underline{\mathbb{R}}^{p \times q}$ partitioned by

$$X = \begin{matrix} & \begin{matrix} t_1 & t_2 \end{matrix} \\ \begin{matrix} s_1 \\ s_2 \end{matrix} & \begin{bmatrix} X_1 & X_{12} \\ X_{21} & X_2 \end{bmatrix} \end{matrix}, \quad Y = \begin{matrix} & \begin{matrix} q_1 & q_2 \end{matrix} \\ \begin{matrix} p_1 \\ p_2 \end{matrix} & \begin{bmatrix} Y_1 & Y_{12} \\ Y_{21} & Y_2 \end{bmatrix} \end{matrix},$$

it follows that

$$\begin{aligned} X \otimes Y &= \begin{bmatrix} X_1 \otimes Y & X_{12} \otimes Y \\ X_{21} \otimes Y & X_2 \otimes Y \end{bmatrix} \\ &= \begin{bmatrix} U_{s_1 \times p} & 0 \\ 0 & U_{s_2 \times p} \end{bmatrix} \begin{bmatrix} Y \otimes X_1 & Y \otimes X_{12} \\ Y \otimes X_{21} & Y \otimes X_2 \end{bmatrix} \begin{bmatrix} U_{q \times t_1} & 0 \\ 0 & U_{q \times t_2} \end{bmatrix} \\ &= \begin{bmatrix} U_{s_1 \times p} & 0 \\ 0 & U_{s_2 \times p} \end{bmatrix} \begin{bmatrix} Y_1 \otimes X_1 & Y_{12} \otimes X_1 & Y_1 \otimes X_{12} & Y_{12} \otimes X_{12} \\ Y_{21} \otimes X_1 & Y_2 \otimes X_1 & Y_{21} \otimes X_{12} & Y_2 \otimes X_{12} \\ Y_1 \otimes X_{21} & Y_{12} \otimes X_{21} & Y_1 \otimes X_2 & Y_{12} \otimes X_2 \\ Y_{21} \otimes X_{21} & Y_2 \otimes X_{21} & Y_{21} \otimes X_2 & Y_2 \otimes X_2 \end{bmatrix} \begin{bmatrix} U_{q \times t_1} & 0 \\ 0 & U_{q \times t_2} \end{bmatrix}, \end{aligned}$$

where U_{ixj} is the permutation matrix defined in Refs. 35 and 36. Since $U_{ixj} = U_{ixj}^{-1}$ (i.e., U_{ixj} is involutory), the stability of (square) $X \otimes Y$ is equivalent to the stability of the above block 4×4 matrix. Hence note that

$$\bar{A}_m = \text{block-diagonal}(\underline{A}, A \oplus A_m, A_m \oplus A, A_m \oplus A_m).$$

If (A_m, B_m, C_m) is such that \bar{A}_m is stable then clearly \underline{A} is also stable and, by elementary properties of the Kronecker sum, A_m is stable. Conversely, if \underline{A} and A_m are stable then \bar{A}_m is stable. The result for \bar{A}_e is obtained analogously noting only that \bar{A}_e is lower block triangular.

5. Appendix C: Proof of Lemma 2.1

(i) From Theorem 6.2.5, p. 123 of Ref. 39, it follows that there exists an $n \times n$ invertible matrix Ψ such that the nonnegative-definite matrices $D_Q \triangleq \Psi^{-1} \hat{Q} \Psi^{-T}$ and $D_P \triangleq \Psi^T \hat{P} \Psi$ are both diagonal. Hence $\underline{C}(\hat{Q}, \hat{P}) \neq \emptyset$. Since $D_Q D_P = \Psi^{-1} \hat{Q} \hat{P} \Psi$ is also diagonal, $\underline{C}(\hat{Q}, \hat{P}) \subset \underline{D}(\hat{Q} \hat{P})$.

(ii) Since $\hat{\hat{Q}} \hat{\hat{P}} = \Psi \Lambda \Psi^{-1}$, where $\Lambda \triangleq D_Q D_P$ is nonnegative diagonal, $\hat{\hat{Q}} \hat{\hat{P}}$ is nonnegative semisimple.

(iii) Since $\hat{\hat{Q}} \hat{\hat{P}}$ is semisimple, it is group invertible. By properties of the group inverse (Ref. 37, p. 124), $\tau^2 = \tau$.

(iv) Note that by means of a basis rearrangement, it can be assumed that Ψ in (i) is such that $\Lambda = \text{diag}(\lambda_1, \dots, \lambda_{n_r}, 0, \dots, 0)$, where $\lambda_i \triangleq (D_Q D_P)_{(i,i)} \neq 0$, $i=1, \dots, n_r$. Hence, since $\Lambda^\# = \text{diag}(\lambda_1^{-1}, \dots, \lambda_{n_r}^{-1}, 0, \dots, 0)$,

$$\tau = \hat{\hat{Q}} \hat{\hat{P}} (\hat{\hat{Q}} \hat{\hat{P}})^\# = \Psi \Lambda \Lambda^\# \Psi^{-1} = \sum_{i=1}^{n_r} \Psi E_i \Psi^{-1}.$$

(v) Since $\rho(\hat{Q}) = \rho(\hat{P}) = \rho(\hat{Q}\hat{P})$, it follows that $(D_{\hat{Q}}^{\wedge})_{(i,i)} \neq 0$, $(D_{\hat{P}}^{\wedge})_{(i,i)} \neq 0$, $i=1, \dots, n_r$. Hence, let $\Lambda_{\hat{Q}}$ and $\Lambda_{\hat{P}}$ denote the upper left positive-diagonal blocks of $D_{\hat{Q}}$ and $D_{\hat{P}}$, respectively, and define

$$\hat{\psi} \triangleq \psi \begin{bmatrix} (\Lambda_{\hat{Q}} \Lambda_{\hat{P}}^{-1})^{\frac{1}{4}} & 0 \\ 0 & I_{n-n_r} \end{bmatrix}.$$

It now follows that

$$\hat{\psi}^{-1} \hat{Q} \hat{\psi}^{-T} = \hat{\psi}^T \hat{P} \hat{\psi}^{-1} = \begin{bmatrix} (\Lambda_{\hat{Q}} \Lambda_{\hat{P}})^{\frac{1}{2}} & 0 \\ 0 & 0 \end{bmatrix} \quad (70)$$

as desired.

(vi) This is an immediate consequence of (70).

(vii) This is an immediate consequence of (70) and (30)

(viii) With ψ as in (iv) and $\Lambda_o \triangleq \text{diag}(\tau_1, \dots, \tau_{n_r})$, it follows that for arbitrary invertible $S \in \mathbb{R}^{n \times n}$,

$$\hat{\hat{Q}}\hat{\hat{P}} = \psi \begin{bmatrix} S \\ 0 \end{bmatrix} (S^{-1} \Lambda_o S) [S^{-1} \ 0] \psi^{-1}$$

and thus (33) and (34) hold with $G = [S^T \ 0]$, $M = S^{-1} \Lambda_o S$ and $\Gamma = [S^{-1} \ 0] \psi^{-1}$.

(ix) The result follows from $S = \bar{M}^{-1} \bar{\Gamma} G^T \bar{M}^{-1}$ with $S^{-1} = M \bar{\Gamma} G^T \bar{M}^{-1}$.

(x) The result is a consequence of (viii) and (ix).

6. Appendix D: Proof of Theorem 2.3.

First note that by arguments similar to those used in Appendix A the dual of (22) given by

$$0 = \tilde{A}_r^T \tilde{P}_r + \tilde{P}_r \tilde{A}_r + \sum_{i=1}^P \tilde{A}_{ri}^T \tilde{P}_r \tilde{A}_{ri} + \tilde{R}_r \quad (71)$$

has a unique, nonnegative-definite solution given explicitly by

$$\tilde{P}_r = \text{vec}^{-1}_{\begin{pmatrix} n_r \\ n_r \end{pmatrix}} (-\tilde{A}_r^{-T} \text{vec } \tilde{R}_r). \quad (72)$$

Define the partitionings

$$\tilde{Q}_r = \begin{matrix} & n & n_r \\ n & \begin{bmatrix} Q_1 & Q_{12} \\ Q_{12}^T & Q_2 \end{bmatrix} \\ n_r & \end{matrix}, \quad \tilde{P}_r = \begin{matrix} & n & n_r \\ n & \begin{bmatrix} P_1 & P_{12} \\ P_{12}^T & P_2 \end{bmatrix} \\ n_r & \end{matrix},$$

where, for notational convenience, we suppress the subscript r . Also define the notation

$$\tilde{\tilde{PQ}} = \begin{bmatrix} Z_1 & Z_{12} \\ Z_{21} & Z_2 \end{bmatrix},$$

where

$$\begin{aligned} Z_1 &\triangleq P_1 Q_1 + P_{12} Q_{12}^T, & Z_{12} &\triangleq P_1 Q_{12} + P_{12} Q_2, \\ Z_{21} &\triangleq P_{12}^T Q_1 + P_2 Q_{12}^T, & Z_2 &\triangleq P_{12}^T Q_{12} + P_2 Q_2, \end{aligned}$$

and let $(\delta_{A_c}, \delta_{B_c}, \delta_{C_c}) \in \mathbb{R}^{n_c \times n_c} \times \mathbb{R}^{n_c \times \ell} \times \mathbb{R}^{m \times n_c}$. We now specialize to the control problem.

Lemma 6.1. Under the assumptions of Theorem 2.3,

$$\begin{aligned}
 J'_c(A_c, B_c, C_c)(\delta_{A_c}, \delta_{B_c}, \delta_{C_c}) &= 2\text{tr}[Z_2^T \delta_{A_c}] \\
 &+ 2\text{tr}[(\hat{V}_2 B_c^T P_2 + CZ_{21}^T + [V_{12}^T + \sum_{i=1}^P C_i Q_i A_i^T] P_{12}) \delta_{B_c}] \\
 &+ 2\text{tr}[(Q_2 C_c^T R_2 + Z_{12}^T B + Q_{12}^T [R_{12} + \sum_{i=1}^P A_i^T P_i B_i]) \delta_{C_c}].
 \end{aligned} \tag{73}$$

Proof. From Lemma 3.7.2, p. 203 of Ref. 38, it follows that the Frechet derivative of the map $W \rightarrow W^{-1}$ is given by $\delta_W \rightarrow -W^{-1} \delta_W W^{-1}$. Also, recall from Refs. 35 and 36 the identities

$$(\text{vec } X)^T \text{vec } Y = \text{tr } X^T Y,$$

$$(X \otimes Y) \text{vec } Z = \text{vec } YZ X^T.$$

Hence, using (23) and noting that \tilde{V}_c and \tilde{R}_c are independent of A_c , we compute

$$\begin{aligned}
 \frac{\partial J_c(A_c, B_c, C_c)}{\partial A_c} \delta_{A_c} &= (\text{vec } \tilde{R}_c)^T \tilde{A}_c^{-1} \left(\frac{\partial \tilde{A}_c}{\partial A_c} \delta_{A_c} \right) \tilde{A}_c^{-1} \text{vec } \tilde{V}_c \\
 &= (\tilde{A}_c^{-T} \text{vec } \tilde{R}_c)^T \left(\begin{bmatrix} 0 & 0 \\ 0 & \delta_{A_c} \end{bmatrix} \oplus \begin{bmatrix} 0 & 0 \\ 0 & \delta_{A_c} \end{bmatrix} \right) (\tilde{A}_c^{-1} \text{vec } \tilde{V}_c) \\
 &= (\text{vec } \tilde{P}_c)^T \left(\begin{bmatrix} 0 & 0 \\ 0 & \delta_{A_c} \end{bmatrix} \oplus I_{n_c}^{-1} + I_{n_c}^{-1} \oplus \begin{bmatrix} 0 & 0 \\ 0 & \delta_{A_c} \end{bmatrix} \right) \text{vec } \tilde{Q}_c \\
 &= 2\text{tr } \tilde{Q}_c^T \tilde{P}_c \begin{bmatrix} 0 & 0 \\ 0 & \delta_{A_c} \end{bmatrix} \\
 &= 2\text{tr } Z_2^T \delta_{A_c}.
 \end{aligned}$$

Furthermore, noting that

$$\text{vec } \tilde{V}_c = \text{vec } \tilde{B}_c \tilde{V}_c^T = (\tilde{B}_c \otimes \tilde{B}_c) \text{vec } \tilde{V},$$

we obtain

$$\begin{aligned} \frac{\partial J_c(A_c, B_c, C_c)}{\partial B_c} \delta_{B_c} &= (\text{vec } \tilde{P}_c)^T \left(\frac{\partial \tilde{A}_c}{\partial B_c} \delta_{B_c} \right) \text{vec } \tilde{Q}_c \\ &\quad - (\text{vec } \tilde{P}_c)^T \left(\frac{\partial (\tilde{B}_c \otimes \tilde{B}_c)}{\partial B_c} \delta_{B_c} \right) \text{vec } \tilde{V} \\ &= (\text{vec } \tilde{P}_c)^T \left(\begin{bmatrix} 0 & 0 \\ \delta_{B_c} C & 0 \end{bmatrix} \oplus \begin{bmatrix} 0 & 0 \\ \delta_{B_c} C & 0 \end{bmatrix} \right. \\ &\quad \left. + \sum_{i=1}^P \tilde{A}_{ci} \otimes \begin{bmatrix} 0 & 0 \\ \delta_{B_c} C_i & 0 \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ \delta_{B_c} C_i & 0 \end{bmatrix} \otimes \tilde{A}_{ci} \right) \text{vec } \tilde{Q}_c \\ &\quad - (\text{vec } \tilde{P}_c)^T \left(\tilde{B}_c \otimes \begin{bmatrix} 0 & 0 \\ 0 & \delta_{B_c} \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ 0 & \delta_{B_c} \end{bmatrix} \otimes \tilde{B}_c \right) \text{vec } \tilde{V} \\ &= 2 \text{tr } \tilde{Q}_c \tilde{P}_c \begin{bmatrix} 0 & 0 \\ \delta_{B_c} C & 0 \end{bmatrix} + 2 \text{tr } \sum_{i=1}^P \tilde{Q}_c \tilde{A}_{ci}^T \tilde{P}_c \begin{bmatrix} 0 & 0 \\ \delta_{B_c} C_i & 0 \end{bmatrix} \\ &\quad + 2 \text{tr } \tilde{V} \tilde{B}_c^T \tilde{P}_c \begin{bmatrix} 0 & 0 \\ 0 & \delta_{B_c} \end{bmatrix} \\ &= 2 \text{tr } C Z_{21}^T \delta_{B_c} + 2 \text{tr } \sum_{i=1}^P (C_i Q_i A_i^T P_{12} + C_i Q_i C_i^T) \\ &\quad + 2 \text{tr} (V_{12}^T P_{12} + V_2 B_c^T P_2) \delta_{B_c} \\ &= 2 \text{tr} (\hat{V}_2 B_c^T P_2 + C Z_{21}^T + [V_{12}^T + \sum_{i=1}^P C_i Q_i A_i^T] P_{12}) \delta_{B_c}. \end{aligned}$$

A similar computation for $(\partial J_c(A_c, B_c, C_c)/\partial C_c)\delta_{C_c}$ yields (73). \square

We can now proceed with the proof of Theorem 2.3. Obviously, (58) is equivalent to

$$0 = Z_2. \quad (74)$$

$$0 = P_2 B_c \hat{V}_2 + Z_{21} C^T + P_{12}^T (V_{12} + \sum_{i=1}^P A_i Q_1 C_i^T). \quad (75)$$

$$0 = \hat{R}_2 C Q_2 + B^T Z_{12} + (R_{12}^T + \sum_{i=1}^P B_i^T P_1 A_i) Q_{12}. \quad (76)$$

Expanding the $n \times n$, $n \times n_c$ and $n_c \times n_c$ blocks of (20) and (71) yields

$$\begin{aligned} 0 = & A Q_1 + Q_1 A^T + V_1 + B C_c Q_{12}^T + Q_{12} (B C_c)^T \\ & + \sum_{i=1}^P [A_i Q_1 A_i^T + B_i C_c Q_{12}^T A_i^T + A_i Q_{12} (B_i C_c)^T + B_i C_c Q_2 (B_i C_c)^T], \end{aligned} \quad (77)$$

$$\begin{aligned} 0 = & A Q_{12} + Q_{12} A_c^T + B C_c Q_2 + Q_1 (B C_c)^T \\ & + \sum_{i=1}^P A_i Q_1 (B C_c)_i^T + V_{12} B_c^T. \end{aligned} \quad (78)$$

$$0 = A_c Q_2 + Q_2 A_c^T + B_c C Q_{12} + Q_{12}^T (B_c C)^T + B_c \hat{V}_2 B_c^T. \quad (79)$$

$$\begin{aligned} 0 = & A^T P_1 + P_1 A + R_1 + (B_c C)^T P_{12}^T + P_{12} B_c C \\ & + \sum_{i=1}^P [A_i^T P_1 A_i + (B_c C_i)^T P_{12}^T A_i + A_i^T P_{12} B_c C_i + (B_c C_i)^T P_2 B_c C_i]. \end{aligned} \quad (80)$$

$$\begin{aligned}
0 &= A^T P_{12} + P_{12} A_c + (B_c C)^T P_2 + P_1 B C_c \\
&\quad + \sum_{i=1}^P A_i^T P_1 B_i C_c + R_{12} C_c.
\end{aligned} \tag{81}$$

$$0 = A_c^T P_2 + P_2 A_c + (B C_c)^T P_{12} + P_{12}^T B C_c + C_c^T \hat{R}_2 C_c. \tag{82}$$

Obviously, $V_2 > 0$ and $R_2 > 0$ imply $\hat{V}_2 > 0$ and $\hat{R}_2 > 0$. Next note that since (A_c, B_c) is controllable and $V_2 > 0$ it follows that $(A_c + B_c C Q_{12} Q_2^{\#}, B_c \hat{V}_2^{\frac{1}{2}})$ is controllable. Using $Q_{12} = Q_{12} Q_2^{\#} Q_2$ (Refs. 39 and 40), (79) can be rewritten as

$$0 = (A_c + B_c C Q_{12} Q_2^{\#}) Q_2 + Q_2 (A_c + B_c C Q_{12} Q_2^{\#})^T + B_c \hat{V}_2 B_c^T. \tag{83}$$

Now, using Lemma 12.2 of Ref. 41 it follows from (83) that Q_2 is positive definite. Similarly, P_2 is positive definite.

Since \hat{R}_2 , \hat{V}_2 , Q_2 and P_2 are invertible, (74)-(76) can be written as

$$-P_2^{-1} P_{12}^T Q_{12} Q_2^{-1} = I_{n_c}. \tag{84}$$

$$B_c = -P_2^{-1} [Z_{21} C^T + P_{12}^T (V_{12} + \sum_{i=1}^P A_i Q_1 C_i^T)] \hat{V}_2^{-1}. \tag{85}$$

$$C_c = -\hat{R}_2^{-1} [B^T Z_{12} + (R_{12}^T + \sum_{i=1}^P B_i^T P_1 A_i) Q_{12}] Q_2^{-1}. \tag{86}$$

Now define new variables

$$Q \triangleq Q_1 - Q_{12} Q_2^{-1} Q_{12}^T, \quad P \triangleq P_1 - P_{12} P_2^{-1} P_{12}^T. \tag{87}$$

$$\hat{Q} \triangleq Q_{12} Q_2^{-1} Q_{12}^T, \quad \hat{P} \triangleq P_{12} P_2^{-1} P_{12}^T. \tag{88}$$

which are $n \times n$ nonnegative-definite matrices. Note that because of (84), (33) and (34) hold with

$$G \triangleq Q_2^{-1} Q_{12}^T, \quad M \triangleq Q_2 P_2, \quad \Gamma \triangleq -P_2^{-1} P_{12}^T, \quad (89)$$

where M is positive semisimple since

$$Q_2 P_2 = Q_2^{\frac{1}{2}} (Q_2^{\frac{1}{2}} P_2 Q_2^{\frac{1}{2}}) Q_2^{\frac{1}{2}}.$$

It is helpful to note the identities

$$\hat{Q} = Q_{12} G = G^T Q_{12}^T = G^T Q_2 G, \quad \hat{P} = -P_{12} \Gamma = -\Gamma^T P_{12}^T = \Gamma^T P_2 \Gamma, \quad (90)$$

$$\tau G^T = G^T, \quad \Gamma \tau = \Gamma, \quad (91)$$

$$\hat{Q} = \tau \hat{Q}, \quad \hat{P} = \hat{P} \tau, \quad (92)$$

$$\hat{\hat{Q}} \hat{\hat{P}} = -Q_{12} P_{12}^T. \quad (93)$$

Using (34) and Sylvester's inequality, it follows that $\rho(G) = \rho(\Gamma) = \rho(Q_{12}) = \rho(P_{12}) = n_c$ which in turn imply (66).

The components of \tilde{Q}_c and \tilde{P}_c can be written in terms of $Q, P, \hat{Q}, \hat{P}, G$ and Γ as

$$Q_1 = Q + \hat{Q}, \quad P_1 = P + \hat{P}, \quad (94)$$

$$Q_{12} = \hat{Q} \Gamma^T, \quad P_{12} = -\hat{P} G^T, \quad (95)$$

$$Q_2 = \Gamma \hat{Q} \Gamma^T, \quad P_2 = G \hat{P} G^T. \quad (96)$$

The gain expressions (60) and (61) can now be seen to be equivalent to (85) and (86). Substituting (94)-(96) into (77)-(82) yields

$$0 = A_Q + Q A^T + V_1 + \sum_{i=1}^P [A_i Q A_i^T + (A_i - B_i \hat{R}_2^{-1} \underline{P}) \hat{Q} (A_i - B_i \hat{R}_2^{-1} \underline{P})^T] + A_P \hat{Q} + \hat{Q} A_P^T. \quad (97)$$

$$0 = [A_P \hat{Q} + \hat{Q} (\Gamma^T A_c G + C^T \hat{V}_2^{-1} Q^T) + Q \hat{V}_2^{-1} Q^T] \Gamma^T, \quad (98)$$

$$0 = \Gamma [(G^T A_c \Gamma + Q \hat{V}_2^{-1} C) \hat{Q} + \hat{Q} (G^T A_c \Gamma + Q \hat{V}_2^{-1} C)^T + Q \hat{V}_2^{-1} Q^T] \Gamma^T, \quad (99)$$

$$0 = A^T P + P A + R_1 + \sum_{i=1}^P [A_i^T P A_i + (A_i - Q \hat{V}_2^{-1} C_i)^T \hat{P} (A_i - Q \hat{V}_2^{-1} C_i)] + A_Q^T \hat{P} + \hat{P} A_Q. \quad (100)$$

$$0 = [A_Q^T \hat{P} + \hat{P} (G^T A_c \Gamma + B \hat{R}_2^{-1} \underline{P}) + \underline{P}^T \hat{R}_2^{-1} \underline{P}] G^T, \quad (101)$$

$$0 = G [(G^T A_c \Gamma + B \hat{R}_2^{-1} \underline{P})^T \hat{P} + \hat{P} (G^T A_c \Gamma + B \hat{R}_2^{-1} \underline{P}) + \underline{P}^T \hat{R}_2^{-1} \underline{P}] G^T. \quad (102)$$

The remaining calculations proceed as follows. Computing either (99)- Γ (98) or (102)- G (101) yields (59). Inserting this expression for A_c into (98), (99), (101) and 102 and computing the (equivalent) equations $G^T(98)^T$, $G^T(99)G$, (101) Γ and $\Gamma^T(102)\Gamma$ it follows that $G^T(99)G = G^T(98)\tau^T$ and $\Gamma^T(102)\Gamma = \tau(101)\Gamma$. Hence, (99) and (102) are superfluous. Furthermore, $G^T(98)^T$ and (101) Γ are equivalent, respectively, to

$$0 = \tau [A_P \hat{Q} + \hat{Q} A_P^T + Q \hat{V}_2^{-1} Q^T], \quad (103)$$

$$0 = [A_Q^T \hat{P} + \hat{P} A_Q + \underline{P}^T \hat{R}_2^{-1} \underline{P}] \tau. \quad (104)$$

Finally, to obtain (62)-(65) note that (64) = (103) + (103) T - (103) τ^T . (62) = (97) - (64). (65) = (104) + (104) T - τ^T (104) and (63) = (100) - (65). \square

References

1. DOYLE, J. C., Guaranteed Margins for LQG Regulators, IEEE Transactions on Automatic Control, Vol. AC-23, pp. 756-757, 1978.
2. SOROKA, E., and SHAKED, U., On the Robustness of LQ Regulators, IEEE Transactions on Automatic Control, Vol. AC-29, pp. 664-665, 1984.
3. WONHAM, W. M., Optimal Stationary Control of a Linear System with State-Dependent Noise, SIAM Journal on Control, Vol. 5, pp. 486-500, 1967.
4. WONHAM, W. M., On a Matrix Riccati Equation of Stochastic Control, SIAM Journal on Control, Vol. 6, pp. 681-697, 1968.
5. KLEINMAN, D., Optimal Stationary Control of Linear Systems With Control-Dependent Noise, IEEE Transactions on Automatic Control, Vol. AC-14, pp. 673-677, 1969.
6. MCLANE, P., Optimal Stochastic Control of Linear Systems with State and Control-Dependent Disturbances, IEEE Transactions on Automatic Control, Vol. AC-16, pp. 793-798, 1971.
7. HAUSSMANN, U., Optimal Stationary Control With State and Control Dependent Noise, SIAM Journal on Control, Vol. 14, pp. 419-444, 1976.
8. BISMUT, J., Linear-Quadratic Optimal Stochastic Control With Random Coefficients, SIAM Journal on Control, Vol. 14, pp. 419-444, 1976.
9. MERRIAM, C. W. III, Automated Design of Control Systems, Gordon and Breach, New York, 1974.
10. GUSTAFSON, D. E., and SPEYER, J. L., Design of Linear Regulators for Nonlinear Systems, Journal of Spacecraft and Rockets, Vol. 12, pp. 351-358, 1975.

11. BERNSTEIN, D. S., and GREELEY, S. W., Robust Controller Synthesis Using the Maximum Entropy Design Equations, IEEE Transactions on Automatic Control, Vol. AC-31, pp. 362-364, 1986.
12. HYLAND, D. C., and BERNSTEIN, D. S., The Optimal Projection Equations for Fixed-Order Dynamic Compensation, IEEE Transactions on Automatic Control, Vol. AC-29, pp. 1034-1037, 1984.
13. HYLAND, D. C., and MADIWALE, A. N., A Stochastic Design Approach for Full-Order Compensation of Structural Systems with Uncertain Parameters, Proceedings of the AIAA Guidance and Control Conference, pp. 324-332, Albuquerque, NM, August 1981.
14. HYLAND, D. C., Mean-Square Optimal, Full-Order Compensation of Structural Systems with Uncertain Parameters, MIT, Lincoln Laboratory TR-626, 1 June 1983.
15. MIL'SHTEIN, G. N., Design of Stabilizing Controller with Incomplete State Data for Linear Stochastic System with Multiplicative Noise, Automation and Remote Control, Vol. 43, pp. 653-659, 1982.
16. STRATONOVICH, R. L., A New Representation for Stochastic Integrals, SIAM Journal on Control, Vol. 4, pp. 362-371, 1966.
17. ARNOLD, L., Stochastic Differential Equations: Theory and Applications, Wiley, New York, 1974.
18. IKEDA, N., and WATANABE, S., Stochastic Differential Equations and Diffusion Processes, North-Holland, Amsterdam, 1981.
19. HYLAND, D. C., Active Control of Large Flexible Spacecraft: A New Design Approach Based on Minimum Information Modelling of Parameter Uncertainties, Proceedings of the Third VPI&SU/AIAA Symposium on Large Space Structures, pp. 631-646, Blacksburg, VA, June 1981.

20. HYLAND, D. C., Maximum Entropy Stochastic Approach to Control Design for Uncertain Structural Systems, Proceedings of the American Control Conference, pp. 680-688, Arlington, VA, June 1982.
21. JAYNES, E. T., Where Do We Stand on Maximum Entropy?, in The Maximum Entropy Formalism, Edited by D. Levine and M. Tribus, MIT Press, pp. 15-118, Cambridge, MA, 1979.
22. WONG, E., and ZAKAI, M., On the Relation Between Ordinary and Stochastic Differential Equations, International Journal of Engineering Science, Vol. 3, pp. 213-229, 1965.
23. ATHANS, M., KU, R. T., and GERSHWIN, S. B., The Uncertainty Threshold Principle: Some Fundamental Limitations of Optimal Decision Making Under Dynamic Uncertainty, IEEE Transactions on Automatic Control, Vol. AC-22, pp. 491-495, 1977.
24. KU, R. J., and ATHANS, M., Further Results on the Uncertainty Threshold Principle, IEEE Transactions on Automatic Control, Vol. AC-22, pp. 866-868, 1977.
25. BERNSTEIN, D. S., AND HYLAND, D. C., The Optimal Projection/Maximum Entropy Approach to Designing Low-Order, Robust Controllers for Flexible Structures, Proceedings of the 24th IEEE Conference on Decision and Control, pp. 745-752, Fort Lauderdale, FL, December 1985.
26. BERNSTEIN, D. S., and GREELEY, S. W., Robust Output-Feedback Stabilization: Deterministic and Stochastic Perspectives, Proceedings of the American Control Conference, pp. 1818-1826, Seattle, Washington, June 1986.

27. BERNSTEIN, D. S., The Optimal Projection Equations for Static and Dynamic Output Feedback: The Singular Case, IEEE Transactions on Automatic Control, to appear.
28. BERNSTEIN, D. S., Robust Stability and Performance via the Extended Optimal Projection Equations for Fixed-Order Dynamic Compensation, submitted.
29. HYLAND, D. C., and BERNSTEIN, D. S., The Majorant Lyapunov Equation: A Nonnegative Matrix Equation for Guaranteed Robust Stability and Performance of Large Scale Systems, IEEE Transactions on Automatic Control, 1987, to appear.
30. HYLAND, D. C., The Optimal Projection Approach to Fixed-Order Compensation: Numerical Methods and Illustrative Results, AIAA 21st Aerospace Sciences Meeting, paper 83-0303, Reno, Nevada, January 1983.
31. HYLAND, D. C., Comparison of Various Controller-Reduction Methods: Suboptimal Versus Optimal Projection, Proceedings of the AIAA Dynamics Specialists Conference, pp. 381-389, Palm Springs, California, May 1984.
32. HYLAND, D. C., and BERNSTEIN, D. S., The Optimal Projection Equations for Model Reduction and the Relationships Among the Methods of Wilson, Skelton and Moore, IEEE Transactions on Automatic Control, Vol. AC-30, pp. 1201-1211, 1985.
33. BERNSTEIN, D. S., and HYLAND, D.C., The Optimal Projection Equations for Reduced-Order State Estimation, IEEE Transactions on Automatic Control, Vol. AC-30, pp. 583-585, 1985.

34. BERNSTEIN, D. S., DAVIS, L. D., GREELEY, S. W., and HYLAND, D. C.,
Numerical Solution of the Optimal Projection/Maximum Entropy Design
Equations for Low-Order, Robust Controller Design, Proceedings of the
24th IEEE Conference on Decision and Control, pp. 1795-1798,
Fort Lauderdale, Florida, December 1985.
35. BREWER, J. W., Kronecker Products and Matrix Calculus in System Theory,
IEEE Transactions on Circuits and Systems, Vol. CAS-25, pp. 772-781,
1978.
36. GRAHAM, A., Kronecker Products and Matrix Calculus, Ellis Horwood,
Chichester, 1981.
37. CAMPBELL, S. L., and MEYER, C. D., Jr., Generalized Inverses of Linear
Transformations, Pitman, London, 1979.
38. FLETT, T. M., Differential Analysis, Cambridge University Press,
Cambridge, 1980.
39. RAO, C. R., and MITRA, S. K., Generalized Inverse of Matrices and Its
Applications, John Wiley and Sons, New York, New York, 1971.
40. ALBERT, A., Conditions for Positive and Nonnegative Definiteness in
Terms of Pseudo Inverse, SIAM Journal on Applied Mathematics, Vol. 17,
pp. 434-440, 1969.
41. KREINDLER, E., and JAMESON A., Conditions for Nonnegativeness of
Partitioned Matrices, IEEE Transactions on Automatic Control,
Vol. AC-17, pp. 147-148, 1972.
42. WONHAM, W. M., Linear Multivariable Control: A Geometric Approach,
Springer-Verlag, New York, New York, 1974.

Robust Controller Synthesis Using the Maximum Entropy Design Equations

DENNIS S. BERNSTEIN AND SCOTT W. GREELEY

Abstract—This note presents an application of the optimality conditions obtained in [1] for dynamic compensation in the presence of state-, control-, and measurement-dependent noise. By solving these equations, which represent a fundamental generalization of standard steady-state LQG theory, a series of increasingly robust control designs is obtained for the example considered in [2].

I. INTRODUCTION

Perhaps the most significant aspect of LQG theory is the *explicit synthesis* of dynamic feedback compensators. In practice, however, LQG suffers from serious defects concerning closed-loop robustness with respect to plant deviations. In particular, LQG controllers may possess arbitrarily small stability margin with respect to parameter variations [2].

One approach to correcting this defect is to rederive the optimality conditions for dynamic compensation in the presence of state-, control-, and measurement-dependent noise [1]. Intuitively speaking, the quadratically optimal feedback controller designed in the presence of such multiplicative disturbances is *automatically desensitized* to actual parameter variations. The optimality conditions now comprise a system of four matrix equations, specifically, two modified Riccati equations and two modified Lyapunov equations, coupled by stochastic effects. This coupling is a graphic reminder of the breakdown of the separation principle in the uncertain plant case. When the uncertainty terms are absent, the equations immediately reduce to the standard pair of separated Riccati equations.

For the special case of full-order compensation in the presence of state-dependent noise only, versions of these equations were discovered independently by Hyland [3]–[5] and Mil'stein [6]. A crucial feature of [1], [3]–[5] is the interpretation of the closed-loop stochastic differential equation according to the Fisk-Stratonovich definition of stochastic integration. For modeling flexible mechanical structures, justification of this interpretation as an appropriate model for *a priori* parameter uncertainty was based upon the maximum entropy principle of Jaynes [1].

A time-varying version of these design equations involving uncorrelated state- and control-dependent noise has been given in [7]. The stochastic interpretation is in the sense of Ito as in [6].

The purpose of the present note is to summarize the *maximum entropy equations* for full-order dynamic feedback compensation. These equations are then applied to Doyle's example [2] to produce a series of quadratically optimal robust controllers. The full optimal projection/maximum entropy design equations, which also account for a constraint on controller order [1], [8], are applied to a more realistic design problem in [9].

II. PROBLEM STATEMENT AND MAXIMUM ENTROPY DESIGN EQUATIONS

To state the optimal dynamic-compensation problem, we require the following notation. Let $x \in \mathbb{R}^n$, $y \in \mathbb{R}^l$, $u \in \mathbb{R}^m$, $A, A_1, \dots, A_p \in \mathbb{R}^{n \times n}$, $B, B_1, \dots, B_p \in \mathbb{R}^{n \times m}$, $C, C_1, \dots, C_p \in \mathbb{R}^{l \times n}$, $R_1 \in \mathbb{R}^{n \times n}$, $R_1 \geq 0$, $R_2 \in \mathbb{R}^{m \times m}$, $R_2 > 0$. Furthermore, let v_1, \dots, v_p be unit-intensity, zero-mean, and mutually uncorrelated white noise processes and let $w_1 \in \mathbb{R}^n$ and $w_2 \in \mathbb{R}^l$ be zero-mean white noise processes with

Manuscript received July 23, 1985; revised October 24, 1985. This work was supported in part by the Air Force Office of Scientific Research under Contract AFOSR F49620-84-C-0015.

The authors are with the Controls Analysis and Synthesis Group, Government Aerospace Systems Division, Harris Corporation, Melbourne, FL 32902.
IEEE Log Number 8407268.

intensities $V_1 \geq 0$ and $V_2 > 0$, respectively, and cross intensity $V_{12} \in \mathbb{R}^{n \times n}$. It is further assumed that v_i , w_i , and $x(0)$ are mutually uncorrelated. We require the technical assumption that, for each i , $B_i \neq 0$ implies $C_i = 0$, i.e., the control- and measurement-dependent noises are uncorrelated.

Optimal Dynamic-Compensation Problem

Given the controlled system

$$\dot{x} = \left(A + \sum_{i=1}^p v_i A_i \right) x + \left(B + \sum_{i=1}^p v_i B_i \right) u + w_1, \quad (2.1)$$

$$y = \left(C + \sum_{i=1}^p v_i C_i \right) x + w_2, \quad (2.2)$$

design an n th-order dynamic compensator

$$\dot{x}_c = A_c x_c + B_c y, \quad (2.3)$$

$$u = C_c x_c \quad (2.4)$$

which minimizes the performance criterion

$$J(A_c, B_c, C_c) = J_x(A_c, B_c, C_c) + J_m(A_c, B_c, C_c) + J_u(A_c, B_c, C_c), \quad (2.5)$$

where

$$J_x(A_c, B_c, C_c) \triangleq \lim_{t \rightarrow \infty} \mathbb{E}[x^T R_1 x],$$

$$J_m(A_c, B_c, C_c) \triangleq \lim_{t \rightarrow \infty} \mathbb{E}[2x^T R_{12} u],$$

$$J_u(A_c, B_c, C_c) \triangleq \lim_{t \rightarrow \infty} \mathbb{E}[u^T R_2 u].$$

To guarantee that J is finite and independent of initial conditions, we restrict (A_c, B_c, C_c) to the (open) set of second-moment-stabilizing triples

$$\mathcal{S} \triangleq \{(A_c, B_c, C_c) : \bar{A}_c \oplus \bar{A}_c + \sum_{i=1}^p \bar{A}_i \otimes \bar{A}_i \text{ is stable}\}$$

where \oplus and \otimes denote Kronecker sum and product and

$$\bar{A}_i \triangleq \begin{bmatrix} A_i & B_i C_c \\ B_i C_c & A_c \end{bmatrix}, \quad \bar{A}_c \triangleq \begin{bmatrix} A_c & B_c C_c \\ B_c C_c & 0 \end{bmatrix},$$

$$A_c \triangleq A + \frac{1}{2} \sum_{i=1}^p A_i^2, \quad B_c \triangleq B + \frac{1}{2} \sum_{i=1}^p A_i B_i, \quad C_c \triangleq C + \frac{1}{2} \sum_{i=1}^p C_i A_i.$$

For convenience in stating the optimality conditions, define the following notation for $Q, P, \hat{Q}, \hat{P} \in \mathbb{R}^{n \times n}$:

$$R_{12} \triangleq R_1 + \sum_{i=1}^p B_i^T (P + \hat{P}) B_i, \quad V_{12} \triangleq V_2 + \sum_{i=1}^p C_i (Q + \hat{Q}) C_i^T,$$

$$Q_c \triangleq Q C_c^T + V_{12} + \sum_{i=1}^p A_i (Q + \hat{Q}) C_i^T, \quad \hat{P}_c \triangleq B_i^T P + R_{12} + \sum_{i=1}^p B_i^T (P + \hat{P}) A_i,$$

$$A_{Qc} \triangleq A_c - Q_c V_{12}^{-1} C_c, \quad A_{Pc} \triangleq A_c - B_c R_{12}^{-1} \hat{P}_c.$$

Theorem 2.1: Suppose $(A_c, B_c, C_c) \in \mathcal{S}$ solves the optimal dynamic-compensation problem. Then there exist $n \times n$ nonnegative-definite matrices Q, P, \hat{Q} , and \hat{P} such that A_c, B_c, C_c are given by

$$A_c = A - B_c R_{12}^{-1} \hat{P}_c - Q_c V_{12}^{-1} C_c, \quad (2.6)$$

$$B_c = Q_c V_{12}^{-1}, \quad (2.7)$$

$$C_c = -R_{12}^{-1} \hat{P}_c, \quad (2.8)$$

and such that the following conditions are satisfied:

$$0 = A_c Q + Q A_c^T + V_1 + \sum_{i=1}^p [A_i Q A_i^T + (A_i - B_c R_{12}^{-1} \hat{P}_c) \hat{Q} (A_i - B_c R_{12}^{-1} \hat{P}_c)^T] - Q_c V_{12}^{-1} Q_c^T, \quad (2.9)$$

$$0 = A_c^T P + P A_c + R_1 + \sum_{i=1}^p [A_i^T P A_i + (A_i - Q_c V_{12}^{-1} C_c)^T \hat{P} (A_i - Q_c V_{12}^{-1} C_c)] - \hat{P}_c^T R_{12}^{-1} \hat{P}_c, \quad (2.10)$$

$$0 = A_{Qc} \hat{Q} + \hat{Q} A_{Qc}^T + Q_c V_{12}^{-1} Q_c^T, \quad (2.11)$$

$$0 = A_{Pc}^T \hat{P} + \hat{P} A_{Pc} + \hat{P}_c^T R_{12}^{-1} \hat{P}_c. \quad (2.12)$$

Remark 2.1: Letting $A_i = 0$, $B_i = 0$ and $C_i = 0$, $i = 1, \dots, p$, it can readily be seen that (2.11) and (2.12) are superfluous and that (2.9) and (2.10) yield the standard separated LQG Riccati equations.

Remark 2.2: Since $R_{12} \geq R_2$, so that $R_{12}^{-1} \leq R_2^{-1}$, it is clear that the control-dependent noise effectively suppresses the regulator gain C_c . Similarly, since $V_{12} \geq V_2$, the measurement-dependent noise suppresses the observer gain B_c . The effect of the terms $A_i Q A_i^T$ is discussed in [1] for modal systems.

III. THE MAXIMUM ENTROPY DESIGN EQUATIONS APPLIED TO DOYLE'S EXAMPLE

As shown in [2], LQG regulators for the example

$$A = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}, \quad B = \begin{bmatrix} 0 \\ b \end{bmatrix}, \quad C = [1 \quad 0],$$

$$V_1 = \sigma \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}, \quad V_{12} = 0, \quad V_2 = 1,$$

$$R_1 = \rho \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}, \quad R_{12} = 0, \quad R_2 = 1,$$

have arbitrarily small stability margin with regard to variations $b + \Delta b$ when σ and ρ are sufficiently large and $b = 1$.

Setting $\sigma = \rho = 60$, it follows that the LQG regulator is only stable for $0.93 \leq b + \Delta b \leq 1.01$. Uncertainty in b can be modeled by setting $p = 1$, $A_1 = 0$, $B_1 = [0 \ b_1]^T$, and $C_1 = 0$. Solving the optimality conditions (2.9)–(2.12) with $b_1 = 0.05, 0.10, 0.15$, and 0.20 yields a series of increasingly robust controller designs with respect to both positive and negative variations Δb (see Table I and Figs. 1 and 2).

CONCLUSION

As demonstrated on the example of [2], the maximum entropy design equations provide a novel method for synthesizing robust feedback controllers. Since the design equations represent a fundamental generalization of standard LQG theory, the approach represents an alternative to LQG-modification techniques. Indeed, these equations are not intended as a device for recovering the gain and phase margins of LQ state-feedback regulators, but rather as a method for designing output-feedback dynamic compensators which are robust with respect to parametric deviations in

TABLE I
DYNAMIC COMPENSATOR GAINS FOR LQG AND MAXIMUM ENTROPY
DESIGNS ($b = 1$, $\sigma = \rho = 60$)

| b_1 | A_c | B_c | C_c | Stability Range of $b + \Delta b$ |
|------------|---|--|-------------------------|--------------------------------------|
| 0 (LQG) | $\begin{bmatrix} -9 & 1 \\ -20 & -9 \end{bmatrix}$ | $\begin{bmatrix} 10 \\ 10 \end{bmatrix}$ | $[-10 \quad -10]$ | (.83, 1.01) |
| .05 | $\begin{bmatrix} -9.233 & 1.0 \\ -20.69 & -7.382 \end{bmatrix}$ | $\begin{bmatrix} 10.23 \\ 12.31 \end{bmatrix}$ | $[-8.382 \quad -8.382]$ | (.88, 1.03) |
| .10 | $\begin{bmatrix} -9.639 & 1.0 \\ -23.27 & -6.318 \end{bmatrix}$ | $\begin{bmatrix} 10.64 \\ 15.95 \end{bmatrix}$ | $[-7.318 \quad -7.318]$ | (.82, 1.08) |
| .15 | $\begin{bmatrix} -10.10 & 1.0 \\ -27.24 & -5.710 \end{bmatrix}$ | $\begin{bmatrix} 11.10 \\ 20.33 \end{bmatrix}$ | $[-6.710 \quad -6.710]$ | (.77, 1.13) |
| .20 | $\begin{bmatrix} -10.69 & 1.0 \\ -32.97 & -5.295 \end{bmatrix}$ | $\begin{bmatrix} 11.69 \\ 26.67 \end{bmatrix}$ | $[-6.295 \quad -6.295]$ | (.72, 1.21) |

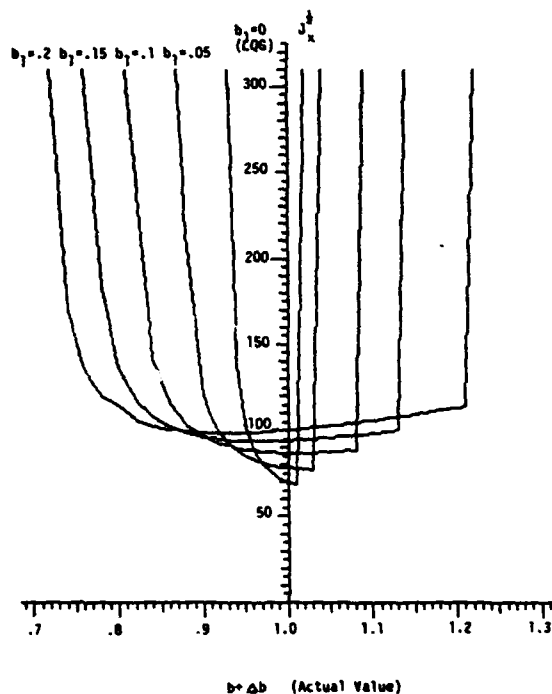


Fig. 1. Robustness of LQG versus maximum entropy designs ($b_1 = 0.05, 0.1, 0.15, 0.2$).

the plant model. As discussed in [10], these are significantly different objectives.

REFERENCES

- [1] D. S. Bernstein and D. C. Hyland, "The optimal projection equations for reduced-order modelling, estimation and control of linear systems with Stratonovich multiplicative white noise," *SIAM J. Contr. Optimiz.*, submitted for publication.
- [2] J. C. Doyle, "Guaranteed margins for LQG regulators," *IEEE Trans. Automat. Contr.*, vol. AC-23, pp. 756-757, 1978.
- [3] D. C. Hyland and A. N. Madiwale, "A stochastic design approach for full-order

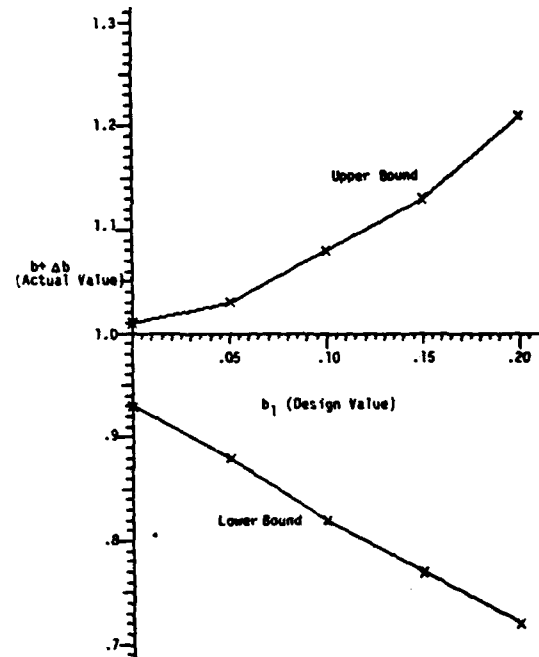


Fig. 2. Stability bounds for LQG and maximum entropy designs.

compensation of structural systems with uncertain parameters," in *Proc. AIAA Guid. Contr. Conf.*, Albuquerque, NM, Aug. 1981, pp. 324-332.

- [4] D. C. Hyland, "Optimality conditions for fixed-order dynamic compensation of flexible spacecraft with uncertain parameters," *AIAA 20th Aerospace Sci. Meet.*, Orlando, FL, Jan. 1982, paper 82-0312.
- [5] D. C. Hyland, "Mean-square optimal, full-order compensation of structural systems with uncertain parameters," Lincoln Lab., Mass. Inst. Technol., Tech. Rep. TR-626, June 1983.
- [6] G. N. Mil'stein, "Design of stabilizing controller with incomplete state data for linear stochastic system with multiplicative noise," *Automat. Remote Contr.*, vol. 43, pp. 653-659, 1982.
- [7] Y. A. Phillis, "Controller design of systems with multiplicative noise," *IEEE Trans. Automat. Contr.*, vol. AC-30, pp. 1017-1019, 1985.
- [8] D. C. Hyland and D. S. Bernstein, "The optimal projection equations for fixed-order dynamic compensation," *IEEE Trans. Automat. Contr.*, vol. AC-29, pp. 1034-1037, 1984.
- [9] L. D. Davis, D. C. Hyland, and D. S. Bernstein, "Application of the maximum entropy design approach to the spacecraft control laboratory experiment (SCOLE)," NASA Langley, Final Report, Jan. 1985.
- [10] D. S. Bernstein, "Robust static and dynamic output-feedback stabilization: Deterministic and stochastic perspectives," submitted for publication.

Robust Static and Dynamic Output-Feedback Stabilization: Deterministic and Stochastic Perspectives

DENNIS S. BERNSTEIN, MEMBER, IEEE

Abstract—Three parallel gaps in robust feedback control theory are examined: sufficiency versus necessity, deterministic versus stochastic uncertainty modeling, and stability versus performance. Deterministic and stochastic output-feedback control problems are considered with both static and dynamic controllers. The static and dynamic robust stabilization problems involve deterministically modeled bounded but unknown measurable time-varying parameter variations, while the static and dynamic stochastic optimal control problems feature state-, control-, and measurement-dependent white noise. General sufficiency conditions for the deterministic problems are obtained using Lyapunov's direct method, while necessary conditions for the stochastic problems are derived as a consequence of minimizing a quadratic performance criterion. The sufficiency tests are then applied to the necessary conditions to determine when solutions of the stochastic optimization problems also solve the deterministic robust stability problems. As an additional application of the deterministic result, the modified Riccati equation approach of Petersen and Hollot is generalized in the static case and extended to dynamic compensation.

I. INTRODUCTION

THE gain and phase margins of full-state-feedback LQ regulators are well known [1], [2]. Although dynamic output-feedback LQG controllers lack such margins [3], considerable effort has been devoted to recovering the full-state-feedback properties [4]–[6]. A crucial point discussed in [7]–[9] is that such margins may be meaningless for guaranteeing robustness with respect to arbitrary plant parameter variations. This was demonstrated by means of a simple example in [7]. In addition, as is well known, the use of singular value bounds to characterize plant uncertainty contributes directly to conservatism with respect to real-valued structured parameter variations.

For the parametric-uncertainty stability robustness problem, there exists a considerable body of literature (see, e.g., [10]–[25]). These results often rely upon Lyapunov's direct method and thus are usually in the form of sufficient conditions. Two factors are often lacking, however: a measure of performance beyond stability and design considerations involving controller effort.¹

Performance and controller effort are, of course, the natural domain of stochastic optimal control via the cost criterion. In addition, parameter uncertainties can be directly incorporated into the stochastic model by means of multiplicative white noise [26]–[40]. Heuristically speaking, the performance of a quadratically optimal feedback controller designed in the presence of such

multiplicative disturbances is desensitized to *actual* constant or time-varying parameter variations. It should be emphasized that the white noise parameter-uncertainty model is not interpreted literally as a physical model. Rather, the multiplicative noise model serves as a device which captures the effect of parameter uncertainty on the second-moment matrix, and hence on the closed-loop performance. From a practical point of view, the multiplicative white noise model is extremely tractable since the second-moment equation is closed and the optimal feedback gains can be given explicitly by closed-form expressions involving solutions of algebraic equations. For example, the necessary conditions derived in [40] for quadratically optimal steady-state fixed-order (i.e., reduced-order) dynamic compensation in the presence of state-, control-, and measurement-dependent white noise involve a coupled system of two modified algebraic Riccati equations and two modified algebraic Lyapunov equations. The coupling is due to both the optimal projection, which enforces the fixed-order constraint [41], and the multiplicative white noise terms. Unfortunately, however, stochastic optimal control is predicated upon stability of the second moment of the state [42]–[47], which may be weaker than deterministic robust asymptotic stability. As a matter of fact, it has been shown, rather surprisingly, that a nominally unstable system can be rendered stochastically stable by multiplicative white noise interpreted in the sense of Stratonovich without actually applying feedback control [48].² Hence, there is no prior guarantee that a second-moment stable optimal controller predicated upon a multiplicative white noise model will provide deterministic robust or even nominal asymptotic stability.

Three parallel gaps can thus be perceived between the above approaches: sufficiency versus necessity, deterministic versus stochastic uncertainty modeling, and stability versus performance. In attempting to bridge these gaps we ask the following question: When is the solution to a stochastic optimal control problem also the solution to a deterministic robust stabilization problem? In the present paper we show that our necessary conditions for stochastic optimality become sufficient conditions for deterministic robustness when we include an exponential weighting factor in the quadratic cost criterion. As shown in [49], the weighting factor $e^{2\alpha t}$ leads to replacement of the closed-loop dynamics matrices

$$A + BKC, \begin{bmatrix} A & BC_c \\ B_c C & A_c \end{bmatrix}$$

for static and dynamic controllers, respectively, by the *shifted* dynamics matrices

$$A + \alpha I_n + BKC, \begin{bmatrix} A + \alpha I_n & BC_c \\ B_c C & A_c + \alpha I_{n_c} \end{bmatrix}.$$

When there are no parametric plant uncertainties, a *right* shift (α

Manuscript received April 23, 1986; revised April 13, 1987. Paper recommended by Associate Editor, D. P. Looze. This work was supported in part by the Air Force Office of Scientific Research under Contracts F49620-86-C-0002 and F49620-86-C-0038.

The author is with the Government Aerospace Systems Division, Melbourne, FL 32902.

IEEE Log Number 8716818.

¹ Several exceptions to these remarks should be noted: necessary and sufficient conditions have been given in [12], [21]; a quadratic cost functional is utilized in [10], [13]; and controller effort is considered in [20].

² This phenomenon does not occur, however, with the Ito interpretation.

> 0) yields a prescribed stability margin, i.e., all closed-loop poles having real part less than $-\alpha$. Unfortunately, a right shift alone does not appear to provide guaranteed stability robustness levels with respect to arbitrary plant variations. Since multiplicative white noise also does not ensure robustness, *the present paper goes beyond previous work by employing the right shift in conjunction with multiplicative white noise to guarantee robust stability over a specified range of deterministic parameter variations*. In particular, we consider perturbations of A , B , and C of the form

$$A + \sum_{i=1}^p \sigma_i(t) A_i, B + \sum_{i=1}^p \sigma_i(t) B_i, C + \sum_{i=1}^p \sigma_i(t) C_i,$$

where A_i , B_i , and C_i denote the *pattern* of parametric uncertainty in A , B , and C and $\sigma_i(\cdot)$ are real-valued Lebesgue measurable functions satisfying

$$|\sigma_i(t)| \leq \delta_i, \quad i = 1, \dots, p. \quad (1.1)$$

In this formulation the patterns A_i , B_i , and C_i are assumed to be known while the deterministically modeled uncertain parameters $\sigma_i(\cdot)$ are known except for the bounds (1.1). Our principal result for both static and dynamic controllers states that a solution of the necessary conditions for stochastic optimal control with exponential weighting and multiplicative white noise provides guaranteed robust asymptotic stability for parametric variations satisfying (1.1) as long as

$$\alpha > \frac{1}{2} \sum_{i=1}^p \delta_i^2 / \gamma_i \quad (1.2)$$

where α is the right shift and γ_i is the intensity of the noise $v_i(t)$ multiplying A_i , B_i , and C_i in the multiplicative white noise formulation of the stochastic optimal control problem. Clearly, the rectangular robust stability region $[-\delta_1, \delta_1] \times \dots \times [-\delta_p, \delta_p]$ in p -dimensional parameter space can be enlarged by increasing either α or $\gamma_1, \dots, \gamma_p$. Note that for given values of α , $\gamma_1, \dots, \gamma_p$ (1.2) does not define a unique robustness region when $p > 1$. The robust stability guarantee holds, however, for simultaneous parameter variations $\sigma_1(\cdot), \dots, \sigma_p(\cdot)$ within each region satisfying (1.2).

The above result is based upon the observation that second-moment stability of a stochastic system with multiplicative disturbances and right-shifted mean dynamics induces a Lyapunov function which guarantees robust stability of a deterministic system subject to time-varying parameter variations. This observation, which appears to have been previously overlooked in the literature, may be utilized in the context of robustness analysis for linear uncertain systems. In the present paper this result is developed within the context of robust controller synthesis to achieve a unified approach to robust, fixed-order controller design consistent with [41].

The derivation of our results is quite simple and is based upon the standard quadratic Lyapunov function

$$V(x) = x^T P x \quad (1.3)$$

where P is given by the *modified* Lyapunov equation

$$0 = \bar{A}^T P + P \bar{A} + 2\alpha P + \sum_{i=1}^p \gamma_i A_i^T P A_i + \bar{R} \quad (1.4)$$

where \bar{A} is the closed-loop dynamics matrix and \bar{R} is a closed-loop weighting matrix. Note that the third and fourth terms in (1.4) correspond to exponential weighting and multiplicative white noise, respectively. The result that $\dot{V}(x) \leq -\gamma \|x\|^2$, with

$\gamma > 0$, follows directly from the inequality

$$\sum_{i=1}^p \sigma_i(t) (A_i^T P + P A_i) \leq \sum_{i=1}^p (\delta_i^2 / \gamma_i) P + \sum_{i=1}^p \gamma_i A_i^T P A_i \quad (1.5)$$

along with (1.2). Inequality (1.5) follows immediately from

$$\sum_{i=1}^p [\gamma_i^{1/2} A_i - (\sigma_i(t) / \gamma_i^{1/2}) I_n]^T P [\gamma_i^{1/2} A_i - (\sigma_i(t) / \gamma_i^{1/2}) I_n] \geq 0 \quad (1.6)$$

and $\sigma_i^2(t) \leq \delta_i^2$.

An alternative approach to guaranteeing robustness of designs predicated upon a multiplicative white noise model is to interpret the stochastic differential equation according to Stratonovich stochastic integration instead of Ito integration [36], [37]. Now the closed-loop dynamics matrices become

$$A_s + B_s K C, \begin{pmatrix} A_s & B_s C_c \\ B_s C_s & A_c \end{pmatrix},$$

where

$$A_s = A + \frac{1}{2} \sum_{i=1}^p \gamma_i A_i^2, B_s = B + \frac{1}{2} \sum_{i=1}^p \gamma_i A_i B_i, C_s = C + \frac{1}{2} \sum_{i=1}^p \gamma_i C_i A_i.$$

The closed-loop Stratonovich *correction* evident in A_s , B_s , and C_s , which can negate the so-called "uncertainty threshold principle" [50], [51], appears to be crucial for designing vibration-suppression controllers for flexible structures [37], [38], [52]. Because of inherent damping, such systems are usually nominally open-loop stable with nondestabilizing uncertainties so that robust stability is less of a challenge than robust performance. Although conditions under which the Stratonovich model yields robust controllers are beyond the scope of the present paper, it should be noted that the differences between the two models are far from trivial. For example, for frequency uncertainties the Stratonovich correction, which corresponds to a variable left shift rather than a uniform right shift, automatically induces a positive-real controller for the high-frequency, poorly modeled flexible modes. Since quadratic Lyapunov functions do not appear to be adequate for guaranteeing the robustness of such designs, the majorant Lyapunov function has been developed [53].

Inasmuch as deterministic robust stability of stochastically optimal controllers is guaranteed by right shift/multiplicative white noise modifications to the closed-loop Lyapunov equation, a natural question which arises is the following: Do there exist *alternative* modifications to the closed-loop Lyapunov equation which guarantee robust stability? One possibility which immediately suggests itself is to replace the bound (1.5) by

$$\sum_{i=1}^p \sigma_i(t) (A_i^T P + P A_i) \leq \sum_{i=1}^p \delta_i (P D_i D_i^T P + E_i^T E_i) \quad (1.7)$$

where $A_i = D_i E_i$. Carrying out full-state-feedback control design with (1.7) leads to an alternative generalization of the standard algebraic regulator Riccati equation. Indeed, a version of this modified Riccati equation has already been developed by Petersen and Hollot [23], [25] as a means for designing robust static and dynamic controllers. A fourth-order aircraft example considered in [25] shows the practical potential of their approach. In the present paper we extend the results of [23], [25] to more general problem formulations encompassing a wider class of parametric uncertainty structures within the context of static output feedback and reduced-order dynamic compensation. Most interestingly, the results we obtain are completely analogous to the static and

dynamic compensation results obtained using a multiplicative white noise model with quadratic cost. This raises the following interesting question: Does there exist an optimization problem whose necessary conditions coincide with the Petersen-Holot-type equations? Indeed, our results were obtained by optimizing over the class of closed-loop Lyapunov equations modified (i.e., robustified) in the sense of Petersen and Holot. Full justification for this technique is developed in [54], [55] where robust performance bounds are obtained.

II. NOTATION AND DEFINITIONS

Note: all matrices have real entries.

| | |
|--|---|
| $\mathbb{R}, \mathbb{R}^{r \times s}, \mathbb{R}^r$ | real numbers, $r \times s$ real matrices, $\mathbb{R}^{r \times 1}$ |
| $I_r, (\cdot)^T$ | $r \times r$ identity matrix, transpose |
| \oplus, \otimes | Kronecker sum, Kronecker product [56] |
| \mathbb{S}^r | $r \times r$ symmetric matrices |
| \mathbb{N}^r | $r \times r$ symmetric nonnegative-definite matrices |
| \mathbb{P}^r | $r \times r$ symmetric positive-definite matrices |
| $Z_1 \geq Z_2$ | $Z_1 - Z_2 \in \mathbb{N}^r, Z_1, Z_2 \in \mathbb{S}^r$ |
| $Z_1 > Z_2$ | $Z_1 - Z_2 \in \mathbb{P}^r, Z_1, Z_2 \in \mathbb{S}^r$ |
| asymptotically stable matrix | matrix with eigenvalues in open left-half plane |
| n, m, l, p, n_c, n_i | positive integers, $i = 1, \dots, p$ |
| \tilde{n}, \tilde{n}_i | $n + n_c, n + n_i, i = 1, \dots, p$ |
| x, u, y, x_c | n, m, l, n_c -dimensional vectors |
| $A, A_i; B, B_i; C, C_i$ | $n \times n$ matrices; $n \times m$ matrices, $l \times n$ matrices; $i = 1, \dots, p$ |
| K | $m \times l$ matrix |
| \tilde{A}, \tilde{A}_i | $A + BKC, A_i + B_iKC, i = 1, \dots, p$ |
| A_c, B_c, C_c | $n_c \times n_c, n_c \times l, m \times n_c$ matrices |
| $\tilde{\tilde{A}}, \tilde{\tilde{A}}_i$ | $\begin{bmatrix} A & BC_c \\ B_cC & A_c \end{bmatrix}, \begin{bmatrix} A_i & B_iC_c \\ B_iC_i & 0 \end{bmatrix}, i = 1, \dots, p$ |
| $\sigma_i(\cdot)$ | Lebesgue measurable function on $[0, \infty), i = 1, \dots, p$ |
| δ_i | positive number, $i = 1, \dots, p$ |
| α | real number |
| $A_\alpha, \tilde{A}_\alpha, \tilde{\tilde{A}}_\alpha$ | $A + \alpha I_n, \tilde{A} + \alpha I_n, A + \alpha I_n$ |
| R_1 | state weighting matrix in \mathbb{N}^n |
| R_2 | control weighting matrix in \mathbb{P}^m |
| R_{12} | $n \times m$ cross weighting matrix such that $R_1 - R_{12}R_2^{-1}R_{12}^T \geq 0$ |
| \tilde{R} | $R_1 + R_{12}KC + (KC)^TR_{12}^T + (KC)^TR_2KC$ |
| $\tilde{\tilde{R}}$ | $\begin{bmatrix} R_1 & R_{12}C_c \\ C_c^TR_{12}^T & C_c^TR_2C_c \end{bmatrix}$ |
| w_i | n -dimensional Wiener process |
| V | incremental covariance of w_i in \mathbb{N}^n |
| w_{1i} | n -dimensional Wiener process |
| w_{2i} | l -dimensional Wiener process |
| V_1 | incremental covariance of w_{1i} in \mathbb{N}^n |
| V_2 | incremental covariance of w_{2i} in \mathbb{P}^l |
| V_{12} | $n \times l$ incremental cross-covariance of w_{1i}, w_{2i} |
| V | $\begin{bmatrix} V_1 & V_{12}B_c^T \\ B_c^TV_{12}^T & B_c^TV_2B_c^T \end{bmatrix}$ |
| v_{ii} | mutually uncorrelated scalar Wiener processes, $i = 1, \dots, p$ |
| γ_i | incremental covariance of $v_{ii}, \gamma_i > 0, i = 1, \dots, p$ |
| \mathbb{E} | expected value |

For the following definitions, let $Q, P, \hat{Q}, \hat{P} \in \mathbb{R}^{n \times n}$:

$$R_{2i} \triangleq R_2 + \sum_{i=1}^p \gamma_i B_i^T P B_i,$$

$$P_i \triangleq B^T P + R_{12}^T + \sum_{i=1}^p \gamma_i B_i^T P A_i,$$

$$R_{2d} \triangleq R_2 + \sum_{i=1}^p \gamma_i B_i^T (P + \hat{P}) B_i,$$

$$P_d \triangleq B^T P + R_{12}^T + \sum_{i=1}^p \gamma_i B_i^T (P + \hat{P}) A_i,$$

$$V_{2d} \triangleq V_2 + \sum_{i=1}^p \gamma_i C_i (Q + \hat{Q}) C_i^T,$$

$$Q_d \triangleq Q C^T + V_{12} + \sum_{i=1}^p \gamma_i A_i (Q + \hat{Q}) C_i^T,$$

$$A_{Pa} \triangleq A_a - B R_{2d}^{-1} P_d,$$

$$A_{Qa} \triangleq A_a - Q_d V_{2d}^{-1} C.$$

III. STATIC OUTPUT FEEDBACK

A. Static Robust Stabilization Problem: Deterministic Sufficiency Theory

Consider the following problem.

Static Robust Stabilization Problem: Determine $K \in \mathbb{R}^{m \times l}$ such that the closed-loop system consisting of the controlled plant

$$\begin{aligned} \dot{x}(t) = & \left(A + \sum_{i=1}^p \sigma_i(t) A_i \right) x(t) \\ & + \left(B + \sum_{i=1}^p \sigma_i(t) B_i \right) u(t), \quad t \in [0, \infty), \end{aligned} \quad (3.1.1)$$

measurements

$$y(t) = Cx(t), \quad (3.1.2)$$

and static output-feedback law

$$u(t) = Ky(t) \quad (3.1.3)$$

is asymptotically stable for all measurable $(\sigma_1, \dots, \sigma_p): [0, \infty) \rightarrow \mathbb{R}^p$ satisfying

$$|\sigma_i(t)| \leq \delta_i, \quad t \in [0, \infty), \quad i = 1, \dots, p. \quad (3.1.4)$$

Remark 3.1.1: The nominal stabilization problem, i.e., the case in which parameter uncertainties are absent, can be recovered by setting $A_i = 0$ and $B_i = 0$. All of the results in this paper can readily be specialized to this case. For brevity, however, the details are omitted.

Remark 3.1.2: The symmetric bounds (3.1.4) are for convenience only. The constraints

$$\rho_i \leq \sigma_i(t) \leq \delta_i, \quad t \in [0, \infty), \quad i = 1, \dots, p \quad (3.1.4)'$$

can be recast in the form (3.1.4) by redefining A and B . Further notational simplification is possible by scaling A_i and B_i so that δ_i

$= 1, i = 1, \dots, p$. For clarity, however, we choose not to do this.

Remark 3.1.3: Standard existence theory guarantees that the absolutely continuous solution $x(\cdot)$ of (3.1.1) exists on $[0, \infty)$ for all $K \in \mathbb{R}^{m \times l}$ and for all measurable $(\sigma_1, \dots, \sigma_p)$ satisfying (3.1.4).

The robustness results are based upon the following easily proved theorem which concerns the construction of a Lyapunov function for robust stability.

Theorem 3.1.1: Let $K \in \mathbb{R}^{m \times l}$ and assume there exist

$$\Phi: \mathbb{P}^n \rightarrow \mathbb{S}^n, \quad (3.1.5)$$

$$P \in \mathbb{P}^n, \quad (3.1.6)$$

$$M_i, N_i \in \mathbb{R}^{n_i \times n}, \quad i = 1, \dots, p \quad (3.1.7)$$

such that

$$M_i^T N_i + N_i^T M_i = \bar{A}_i^T P + P \bar{A}_i, \quad i = 1, \dots, p, \quad (3.1.8)$$

$$0 = \bar{A}^T P + P \bar{A} + \Phi(P), \quad (3.1.9)$$

$$\sum_{i=1}^p \delta_i (M_i^T M_i + N_i^T N_i) < \Phi(P). \quad (3.1.10)$$

Then K solves the static robust stabilization problem.

Proof: Using (3.1.8), compute for $t \in [0, \infty)$ and $i = 1, \dots, p$,

$$\begin{aligned} 0 &\leq [\delta_i^{1/2} M_i - \delta_i^{-1/2} \sigma_i(t) N_i]^T [\delta_i^{1/2} M_i - \delta_i^{-1/2} \sigma_i(t) N_i] \\ &= \delta_i M_i^T M_i + \delta_i^{-1} \sigma_i^2(t) N_i^T N_i - \sigma_i(t) [M_i^T N_i + N_i^T M_i] \\ &\leq \delta_i [M_i^T M_i + N_i^T N_i] - \sigma_i(t) [\bar{A}_i^T P + P \bar{A}_i]. \end{aligned}$$

Thus,

$$\sigma_i(t) [\bar{A}_i^T P + P \bar{A}_i] \leq \delta_i (M_i^T M_i + N_i^T N_i), \quad t \in [0, \infty), i = 1, \dots, p. \quad (3.1.11)$$

Defining the Lyapunov function

$$V(x) \triangleq x^T P x$$

its derivative along solutions $x(t)$ of (3.1.1)–(3.1.3) is given by

$$\begin{aligned} \dot{V}(x(t)) &= x(t)^T P \dot{x}(t) + x(t)^T \dot{P} x(t) \\ &= x(t)^T [\bar{A}^T P + P \bar{A}] x(t) \\ &\quad + x(t)^T \left[\sum_{i=1}^p \sigma_i(t) (\bar{A}_i^T P + P \bar{A}_i) \right] x(t). \end{aligned}$$

Using (3.1.9) and (3.1.11) yields

$$\dot{V}(x(t)) \leq -x(t)^T \left[\Phi(P) - \sum_{i=1}^p \delta_i (M_i^T M_i + N_i^T N_i) \right] x(t).$$

By (3.1.10), there exists $\gamma > 0$ such that $\dot{V}(x(t)) \leq -\gamma \|x(t)\|^2$, $t \in [0, \infty)$, as required. \square

Remark 3.1.4: As will be seen in later sections, this result is applied by choosing M_i and N_i to satisfy

$$N_i^T M_i = P \bar{A}_i, \quad (3.1.12)$$

so that (3.1.8) holds. The bound Φ is then constructed to satisfy (3.1.10).

B. Static Optimal Control Problem: Stochastic Necessity Theory

We now turn to the static optimal control problem with state-dependent and control-dependent white noise and exponentially weighted quadratic cost.

Static Optimal Control Problem: Determine $K \in \mathbb{R}^{m \times l}$ such that, for the closed-loop system consisting of the controlled plant

$$\begin{aligned} dx_t &= A x_t dt + \sum_{i=1}^p A_i x_t dv_{it} + B u_t dt \\ &\quad + \sum_{i=1}^p B_i u_t dv_{it} + e^{-\alpha t} dw_t, \quad t \in [0, \infty), \end{aligned} \quad (3.2.1)$$

measurements

$$y_t = C x_t, \quad (3.2.2)$$

and output-feedback law

$$u_t = K y_t, \quad (3.2.3)$$

the performance criterion

$$J_\alpha(K) \triangleq \lim_{t \rightarrow \infty} \mathbb{E} e^{2\alpha t} [x_t^T R_1 x_t + 2x_t^T R_{12} u_t + u_t^T R_2 u_t] \quad (3.2.4)$$

is minimized.

Remark 3.2.1: The exponential time weighting of the disturbance noise in (3.2.1) is required to balance the exponential weighting in the cost (3.2.4). It has no physical significance as such.

To develop necessary conditions for this problem, K must be restricted to the set of second-moment-stabilizing gains

$$\begin{aligned} \mathcal{S}_\alpha \triangleq \left\{ K \in \mathbb{R}^{m \times l} : \bar{A}_\alpha \oplus \bar{A}_\alpha \right. \\ \left. + \sum_{i=1}^p \gamma_i \bar{A}_i \otimes \bar{A}_i \text{ is asymptotically stable} \right\}. \end{aligned}$$

For the shifted plant dynamics. The requirement $K \in \mathcal{S}_\alpha$ implies the existence of the steady-state nonnegative-definite covariance $Q \triangleq \lim_{t \rightarrow \infty} \mathbb{E} [e^{2\alpha t} x_t x_t^T]$. Furthermore, Q is the unique solution to the modified Lyapunov equation

$$0 = \bar{A}_\alpha Q + Q \bar{A}_\alpha^T + \sum_{i=1}^p \gamma_i \bar{A}_i Q \bar{A}_i^T + V. \quad (3.2.5)$$

An additional technical assumption is that K be confined to the set

$$\mathcal{S}_\alpha^+ \triangleq \{ K \in \mathcal{S}_\alpha : C Q C^T > 0, \text{ where } Q \text{ satisfies (3.2.5)} \}.$$

The positive definiteness condition holds, for example, when Q is positive definite and C has full row rank.

Theorem 3.2.1: Suppose $K \in \mathcal{S}_\alpha^+$ solves the static optimal control problem. Then there exist $P, \bar{Q} \in \mathbb{N}^n$ such that K is given by

$$K = -R_2^{-1} P_2 Q C^T (C Q C^T)^{-1} \quad (3.2.6)$$

and such that P and Q satisfy

$$\begin{aligned} 0 &= A^T P + P A + \sum_{i=1}^p \gamma_i A_i^T P A_i + R_1 \\ &\quad - P^T R_2^{-1} P_2 + \bar{P}^T P_1^T R_2^{-1} P_1 \bar{P}, \end{aligned} \quad (3.2.7)$$

$$0 = (A_n - BR_n^{-1}P_n\hat{r})Q + Q(A_n - BR_n^{-1}P_n\hat{r})^T + \sum_{i=1}^p \gamma_i (A_i - B_i R_n^{-1}P_n\hat{r})Q(A_i - B_i R_n^{-1}P_n\hat{r})^T + V \quad (3.2.8)$$

where

$$\hat{r} \triangleq QC^T(CQC^T)^{-1}C, \hat{r}_1 \triangleq I_n - \hat{r}. \quad (3.2.9)$$

Proof: First note that from [49] it follows that the exponential factors in (3.2.1) and (3.2.4) are equivalent to replacing A by A_n . From [57, Theorem 8.5.5, p. 142], it follows that the state covariance $Q(t) \triangleq \mathbb{E}[e^{2\alpha t}x_t x_t^T]$ satisfies

$$\dot{Q}(t) = A_n Q(t) + Q(t)A_n^T + \sum_{i=1}^p \gamma_i A_i Q(t)A_i^T + V.$$

Since $K \in \mathcal{S}_\alpha$, $Q \triangleq \lim_{t \rightarrow \infty} Q(t)$ exists, is nonnegative definite, and is the unique solution to (3.2.5). Note that

$$J_\alpha(K) = \text{tr } Q\bar{R}.$$

Now define the Lagrangian

$$L(Q, K) \triangleq \text{tr} \left[\lambda Q\bar{R} + (\bar{A}_n Q + Q\bar{A}_n^T + \sum_{i=1}^p \gamma_i \bar{A}_i Q\bar{A}_i^T + V)P \right]$$

with multipliers $\lambda \geq 0$ and $P \in \mathbb{R}^{n \times n}$, and compute

$$\partial L / \partial Q = \bar{A}_n^T P + P\bar{A}_n + \sum_{i=1}^p \gamma_i \bar{A}_i^T P\bar{A}_i + \lambda \bar{R}.$$

Setting $\partial L / \partial Q = 0$, $\lambda = 0$ implies $P = 0$ since $K \in \mathcal{S}_\alpha$. Hence, without loss of generality set $\lambda = 1$ so that

$$0 = \bar{A}_n^T P + P\bar{A}_n + \sum_{i=1}^p \gamma_i \bar{A}_i^T P\bar{A}_i + \bar{R}. \quad (3.2.10)$$

Since P is the (unique) steady-state covariance of the dual system, it is nonnegative definite. Also, since \mathcal{S}_α^+ is open, evaluating $\partial L / \partial K = 0$ yields

$$0 = R_n K C Q C^T + P_n Q C^T.$$

Since $K \in \mathcal{S}_\alpha^+$, CQC^T is invertible, and hence (3.2.6) holds. Finally, (3.2.8) is equivalent to (3.2.5) and, performing some algebraic manipulation, (3.2.7) is equivalent to (3.2.10). \square

Remark 3.2.2: Because $R_n \geq R_1$, and thus $R_n^{-1} \leq R_1^{-1}$, the gain K in (3.2.6) leads to a net decrease in controller authority due to the control-dependent noise. For problems which are open-loop stable and which remain stable under plant variations, this is an intuitively expected consequence of parameter uncertainty.

Remark 3.2.3: Theorem 3.2.1 generalizes and unifies several previous results. In particular, the noise correlation and output feedback constraint constitute generalizations of [26]–[33]. Furthermore, because of the presence of multiplicative noise, the results of [58], [59] are extended. The role of the oblique projection \hat{r} has been discussed in [58], [59]. Connections with the oblique projection τ arising in the dynamic-compensation problem [14] are discussed in [60].

C. Sufficiency Meets Necessity: A Marriage of the Deterministic and Stochastic

We now answer our main question: Can a feedback law predicated on a stochastic multiplicative noise model provide guaranteed deterministic robust asymptotic stability? The answer is "yes" provided the exponential is of sufficient magnitude.

Theorem 3.3.1: Suppose there exists $P \in \mathbb{P}^n$ and $Q \in \mathbb{N}^n$ satisfying $CQC^T > 0$, (3.2.7)–(3.2.9) and

$$0 < (2\alpha - \sum_{i=1}^p \delta_i^2 / \gamma_i)P + \bar{R} \quad (3.3.1)$$

where K in \bar{R} is given by (3.2.6). Then K solves the static robust stabilization problem.

Proof: In Theorem 3.1.1 define

$$\Phi(P) = 2\alpha P + \sum_{i=1}^p \gamma_i \bar{A}_i^T P \bar{A}_i + \bar{R},$$

$$M_i = (\gamma_i / \delta_i)^{1/2} P^{1/2} \bar{A}_i, N_i = (\delta_i / \gamma_i)^{1/2} P^{1/2}.$$

Note that (3.1.5)–(3.1.8) hold. Furthermore, because of the equivalence of (3.2.10) and (3.2.7), it follows that (3.1.9) is equivalent to (3.2.7). Finally, (3.1.10) is a consequence of (3.3.1). \square

Remark 3.3.1: Note that (3.2.7)–(3.2.9) serve to construct a Lyapunov function guaranteeing robust stability. Hence, it is not necessary to actually verify that $K \in \mathcal{S}_\alpha^+$.

By strengthening (3.3.1) the following simplification is immediate.

Corollary 3.3.1: Suppose there exists $P \in \mathbb{P}^n$ and $Q \in \mathbb{N}^n$ satisfying $CQC^T > 0$, (3.2.7)–(3.2.9) and

$$\alpha > \frac{1}{2} \sum_{i=1}^p \sigma_i^2 / \gamma_i. \quad (3.3.2)$$

Then K given by (3.2.6) solves the static robust stabilization problem.

It is interesting to note that the feedback gain given by Corollary 3.3.1 may be an extremal, i.e., local minimum, local maximum, etc., and not necessarily a solution of the static optimal control problem. The result is valid, however, for all extremals of the optimization problem. By specializing Corollary 3.3.1 to a solution, i.e., global minimum, of the optimal control problem, we can bridge the gap between sufficiency and necessity.

Corollary 3.3.2: Suppose $K \in \mathcal{S}_\alpha^+$ solves the static optimal control problem where α satisfies (3.3.1), and suppose that the corresponding solution P of (3.2.7) is positive definite. Then K also solves the static robust stabilization problem.

IV. DYNAMIC OUTPUT FEEDBACK

A. Dynamic Robust Stabilization Problem: Deterministic Sufficiency Theory

Consider the following problem.

Dynamic Robust Stabilization Problem: Determine (A_c, B_c, C_c) such that the closed-loop system consisting of the controlled plant (3.1.1), measurements

$$y(t) = \left(C + \sum_{i=1}^p \sigma_i(t) C_i \right) x(t) \quad (4.1.1)$$

and dynamic output-feedback law

$$\dot{x}_c(t) = A_c x_c(t) + B_c y(t), \quad (4.1.2)$$

$$u(t) = C_c x_c(t) \quad (4.1.3)$$

is asymptotically stable for all measurable $(\sigma_1, \dots, \sigma_p): [0, \infty) \rightarrow \mathbb{R}^p$ satisfying (3.1.4).

Remark 4.1.1: Note that the problem statement places no restriction on the order n_c of the dynamic compensator. Also, we now permit uncertainties in the observation matrix C by including perturbations $\sigma_i(t)C_i$ in (4.1.1).

The following result is completely analogous to Theorem 3.1.1.
Theorem 4.1.1: Given (A_c, B_c, C_c) assume there exist

$$\Phi: \mathbb{P}^n \rightarrow \mathbb{S}^n, \quad (4.1.4)$$

$$\tilde{P} \in \mathbb{P}^n, \quad (4.1.5)$$

$$\tilde{M}_i, \tilde{N}_i \in \mathbb{R}^{n_i \times n}, \quad i=1, \dots, p \quad (4.1.6)$$

such that

$$\tilde{M}_i^T \tilde{N}_i + \tilde{N}_i^T \tilde{M}_i = \tilde{A}_i^T \tilde{P} + \tilde{P} \tilde{A}_i, \quad i=1, \dots, p, \quad (4.1.7)$$

$$0 = \tilde{A}^T \tilde{P} + \tilde{P} \tilde{A} + \Phi(\tilde{P}), \quad (4.1.8)$$

$$\sum_{i=1}^p \delta_i (\tilde{M}_i^T \tilde{M}_i + \tilde{N}_i^T \tilde{N}_i) < \Phi(\tilde{P}). \quad (4.1.9)$$

Then (A_c, B_c, C_c) solves the dynamic robust-stabilization problem.

B. Dynamic Optimal Control Problem: Stochastic Necessity Theory

We now consider the dynamic optimal control problem with state-, control-, and measurement-dependent white noise and exponentially weighted quadratic cost. The optimization is performed over the class of dynamic compensators of fixed order $n_c \leq n$.

Dynamic Optimal Control Problem: Determine (A_c, B_c, C_c) such that, for the closed-loop system consisting of the controlled plant

$$\begin{aligned} dx_t = & A x_t dt + \sum_{i=1}^p A_i x_t dv_{it} + B u_t dt \\ & + \sum_{i=1}^p B_i u_t dv_{it} + e^{-\alpha t} dw_{1t}, \quad t \in [0, \infty) \end{aligned} \quad (4.2.1)$$

measurements

$$dy_t = C x_t dt + \sum_{i=1}^p C_i x_t dv_{it} + e^{-\alpha t} dw_{2t} \quad (4.2.2)$$

and dynamic output-feedback law

$$dx_{ct} = A_c x_{ct} dt + B_c dy_t, \quad (4.2.3)$$

$$u_t = C_c x_{ct} \quad (4.2.4)$$

the performance criterion

$$\begin{aligned} J_\alpha(A_c, B_c, C_c) \triangleq & \lim_{t \rightarrow \infty} \mathbb{E} e^{2\alpha t} \{x_t^T R_1 x_t \\ & + 2x_t^T R_{12} u_t + u_t^T R_2 u_t\} \end{aligned} \quad (4.2.5)$$

is minimized.

To develop necessary conditions we restrict (A_c, B_c, C_c) to the set

$$\mathcal{D}_\alpha \triangleq \left\{ (A_c, B_c, C_c): \tilde{A}_c \oplus \tilde{A}_\alpha + \sum_{i=1}^p \gamma_i \tilde{A}_i \otimes \tilde{A}_i \text{ is asymptotically stable and } (A_c, B_c, C_c) \text{ is minimal} \right\}$$

and invoke the technical assumption

$$[B_i \neq 0 = C_i = 0], \quad i=1, \dots, p. \quad (4.2.6)$$

The following lemma will be needed.

Lemma 4.2.1: If $\tilde{Q}, \tilde{P} \in \mathbb{M}^n$ and $\text{rank } \tilde{Q}\tilde{P} = n_c$, then there exist $G, \Gamma \in \mathbb{R}^{n_c \times n}$ and invertible $M \in \mathbb{R}^{n_c \times n_c}$ such that

$$\tilde{Q}\tilde{P} = G^T M \Gamma, \quad (4.2.7)$$

$$\Gamma G^T = I_{n_c}. \quad (4.2.8)$$

Furthermore, G, M , and Γ are unique modulo a change of basis in $\mathbb{R}^{n_c \times n_c}$.

Proof: The result is an immediate consequence of [61, Theorem 6.2.5, p. 123]. \square

Note that because of (4.2.8), the $n \times n$ matrix $\tau = G^T \Gamma$ is idempotent, i.e., $\tau^2 = \tau$. Since τ is not necessarily symmetric, it is an oblique projection.

Theorem 4.2.1: Suppose (4.2.6) holds and $(A_c, B_c, C_c) \in \mathcal{D}_\alpha$ solves the dynamic optimal control problem. Then there exist $\tilde{Q}, \tilde{P}, \tilde{Q}, \tilde{P} \in \mathbb{M}^n$ such that A_c, B_c, C_c are given by

$$A_c = \Gamma(A - BR_{2d}^{-1}P_d - Q_d V_{2d}^{-1}C)G^T, \quad (4.2.9)$$

$$B_c = \Gamma Q_d V_{2d}^{-1}, \quad (4.2.10)$$

$$C_c = -R_{2d}^{-1}P_d G^T \quad (4.2.11)$$

and such that $\tilde{Q}, \tilde{P}, \tilde{Q}, \tilde{P}$ satisfy

$$\begin{aligned} 0 = & A_\alpha \tilde{Q} + \tilde{Q} A_\alpha^T + V_1 + \sum_{i=1}^p \gamma_i \left[A_i \tilde{Q} A_i^T + (A_i - B_i R_{2d}^{-1} P_d) \tilde{Q} (A_i \right. \\ & \left. - B_i R_{2d}^{-1} P_d)^T \right] - Q_d V_{2d}^{-1} Q_d^T + \tau_\perp Q_d V_{2d}^{-1} Q_d^T \tau_\perp^T, \end{aligned} \quad (4.2.12)$$

$$\begin{aligned} 0 = & A_\alpha^T \tilde{P} + \tilde{P} A_\alpha + R_1 + \sum_{i=1}^p \gamma_i \left[A_i^T \tilde{P} A_i + (A_i - Q_d V_{2d}^{-1} C_i)^T \tilde{P} (A_i \right. \\ & \left. - Q_d V_{2d}^{-1} C_i) \right] - P_d^T R_{2d}^{-1} P_d + \tau_\perp^T P_d^T R_{2d}^{-1} P_d \tau_\perp, \end{aligned} \quad (4.2.13)$$

$$0 = A_{P\alpha} \tilde{Q} + \tilde{Q} A_{P\alpha}^T + Q_d V_{2d}^{-1} Q_d^T - \tau_\perp Q_d V_{2d}^{-1} Q_d^T \tau_\perp^T, \quad (4.2.14)$$

$$0 = A_{Q\alpha}^T \tilde{P} + \tilde{P} A_{Q\alpha} + P_d^T R_{2d}^{-1} P_d - \tau_\perp^T P_d^T R_{2d}^{-1} P_d \tau_\perp, \quad (4.2.15)$$

$$\text{rank } \tilde{Q} = \text{rank } \tilde{P} = \text{rank } \tilde{Q}\tilde{P} = n_c \quad (4.2.16)$$

(4.2.7) and (4.2.8), where

$$\tau \triangleq G^T \Gamma, \quad \tau_\perp \triangleq I_n - \tau. \quad (4.2.17)$$

Proof: As in the proof of Theorem 3.2.1 we note that the exponential factors in (4.2.1), (4.2.2), and (4.2.5) are equivalent to replacing A and A_c by $A + \alpha I_n$ and $A_c + \alpha I_{n_c}$, respectively. Theorem 4.2.1 now follows immediately from [40, Theorem 2.3]. It need only be noted that (4.2.9) follows from

$$A_c + \alpha I_{n_c} = \Gamma(A_\alpha - BR_{2d}^{-1}P_d - Q_d V_{2d}^{-1}C)G^T$$

and the fact that $\Gamma A_\alpha G^T = \Gamma A G^T + \alpha I_{n_c}$ because of (4.2.8).

C. Sufficiency Meets Necessity: The Dynamic Case

We now bridge the gap between sufficiency and necessity for dynamic controllers.

Theorem 4.3.1: Assume (4.2.6) holds and suppose there exists $\tilde{Q}, \tilde{P}, \tilde{Q}, \tilde{P} \in \mathbb{M}^n$ satisfying (4.2.12)–(4.2.17), (4.2.7), (4.2.8),

$$\tilde{P} \triangleq \begin{bmatrix} P + \tilde{P} & -\tilde{P} G^T \\ -G \tilde{P} & G \tilde{P} G^T \end{bmatrix} > 0 \quad (4.3.1)$$

and

$$0 < \left(2\alpha - \sum_{i=1}^p \delta_i^2 / \gamma_i \right) \bar{P} + \bar{R}. \quad (4.3.2)$$

Then (A_c, B_c, C_c) given by (4.2.9)–(4.2.11) solves the dynamic robust stabilization problem.

Remark 4.3.1: Note that \bar{P} is always at least nonnegative definite since

$$\bar{P} = \begin{bmatrix} P & 0 \\ 0 & 0 \end{bmatrix} + \begin{bmatrix} \bar{P}^{1/2} \\ -G\bar{P}^{1/2} \end{bmatrix} \begin{bmatrix} \bar{P}^{1/2} \\ -G\bar{P}^{1/2} \end{bmatrix}^T \geq 0. \quad (4.3.3)$$

Proof of Theorem 4.3.1: As shown in [40], (4.2.6)–(4.2.17) are equivalent to

$$0 = \bar{A}_\alpha^T \bar{P} + \bar{P} \bar{A}_\alpha + \sum_{i=1}^p \gamma_i \bar{A}_i^T \bar{P} \bar{A}_i + \bar{R}, \quad (4.3.4)$$

$$0 = \bar{A}_\alpha \bar{Q} + \bar{Q} \bar{A}_\alpha^T + \sum_{i=1}^p \gamma_i \bar{A}_i \bar{Q} \bar{A}_i^T + \bar{V} \quad (4.3.5)$$

where

$$\bar{Q} = \begin{bmatrix} Q + \bar{Q} & \bar{Q} \Gamma^T \\ \Gamma \bar{Q} & \Gamma \bar{Q} \Gamma^T \end{bmatrix}.$$

The result now follows from Theorem 4.1.1 as in the proof of Theorem 3.3.1. \square

Remark 4.3.2: As in Corollary 3.3.1 the inequality (4.3.2) can be replaced by the stronger condition (3.3.2).

V. THE PETERSEN–HOLLOT APPROACH TO ROBUST STABILIZATION

A. Static Output Feedback

The deterministic Riccati equation approach of Petersen and Hollot is based upon factoring A_i and B_i as

$$A_i = D_i E_i, \quad B_i = D_i F_i, \quad i = 1, \dots, p \quad (5.1.1)$$

where $D_i \in \mathbb{R}^{n \times n_i}$, $E_i \in \mathbb{R}^{n_i \times n}$, and $F_i \in \mathbb{R}^{n_i \times m}$. Obviously, such a factorization may not be unique, and the nonuniqueness is an element of the sufficiency test. To state the sufficiency condition we shall require the notation

$$R_{2a} \triangleq R_2 + \sum_{i=1}^p \delta_i F_i^T F_i, \quad P_a \triangleq B^T P + R_{12} + \sum_{i=1}^p \delta_i F_i^T E_i,$$

$$Q_a \triangleq Q C^T + V_{12}, \quad D \triangleq \sum_{i=1}^p \delta_i D_i D_i^T, \quad E \triangleq \sum_{i=1}^p \delta_i E_i^T E_i.$$

Theorem 5.1.1: Assume there exist $P \in \mathbb{P}^n$ and $Q \in \mathbb{N}^n$ satisfying $Q C Q^T > 0$,

$$0 = A^T P + P A + P D P + R_1 + E - P_a^T R_{2a}^{-1} P_a + \hat{r}^T P_a^T R_{2a}^{-1} P_a \hat{r}, \quad (5.1.2)$$

$$0 = (A + D P - B R_{2a}^{-1} P_a \hat{r}) Q + Q (A + D P - B R_{2a}^{-1} P_a \hat{r})^T + V \quad (5.1.3)$$

and

$$\bar{R} > 0 \quad (5.1.4)$$

where \hat{r} is given by (3.2.9) and K in \bar{R} is given by

$$K = -R_{2a}^{-1} P_a Q C^T (C Q C^T)^{-1}. \quad (5.1.5)$$

Proof: In Theorem 3.1.1 define

$$\Phi(P) = \bar{R} + \sum_{i=1}^p \delta_i [(E_i + F_i K C)^T (E_i + F_i K C) + P D_i D_i^T P],$$

$$M_i = E_i + F_i K C, \quad N_i = D_i^T P$$

so that (3.1.5)–(3.1.8) hold. Note that (5.1.2) and (5.1.3) are equivalent to (3.1.9). Finally, because of (5.1.4), (3.1.10) is satisfied. \square

Remark 5.1.1: Theorem 5.1.1 generalizes [25, Theorem 3.3] in four distinct ways. First, in [25], the uncertainties A and B are assumed to be independent. In our setup this corresponds to the additional condition that $A_i \neq 0$ implies $B_i = 0$. In this case (5.1.1) is satisfied with either $E_i = 0$ or $F_i = 0$ for each i . Second, in [25], A_i and B_i are confined to have unit rank. This would be the case if we required $n_i = 1$ in (5.1.1). Third, in [25], $R_{12} = 0$. And, fourth, the results of [25] apply only to the state-feedback case, i.e., $C = I_n$. In this case (5.1.3) is superfluous and $\hat{r} = I_n$.

B. Dynamic Output Feedback

We now extend the Petersen–Hollot approach to reduced-order dynamic compensation. Our only constraint is that we do not permit uncertainty in the observation matrix. Define

$$A_p \triangleq A - B R_{2a}^{-1} P_a, \quad A_Q \triangleq A - Q_a V_2^{-1} C.$$

Theorem 5.2.1: Assume $C_1 = \dots = C_p = 0$ and suppose there exist $Q, P, \bar{Q}, \bar{P} \in \mathbb{N}^n$ satisfying

$$0 = [A + D(P + \bar{P})] Q + Q [A + D(P + \bar{P})]^T + V_1 - Q_a V_2^{-1} Q_a^T + \tau_1 Q_a V_2^{-1} Q_a^T \tau_1^T, \quad (5.2.1)$$

$$0 = A^T P + P A + R_1 + E + P D P - P_a^T R_{2a}^{-1} P_a + \tau_1^T P_a^T R_{2a}^{-1} P_a \tau_1, \quad (5.2.2)$$

$$0 = (A_p + D P) \bar{Q} + \bar{Q} (A_p + D P)^T + Q_a V_2^{-1} Q_a^T - \tau_1 Q_a V_2^{-1} Q_a^T \tau_1^T, \quad (5.2.3)$$

$$0 = (A_Q + D P)^T \bar{P} + \bar{P} (A_Q + D P) + \bar{P} D \bar{P} + P_a^T R_{2a}^{-1} P_a - \tau_1^T P_a^T R_{2a}^{-1} P_a \tau_1, \quad (5.2.4)$$

(4.2.7), (4.2.8), (4.2.16), (4.3.1), and

$$\begin{bmatrix} R_1 & -R_{12} R_{2a}^{-1} P_a G^T \\ -G P_a^T R_{2a}^{-1} R_{12}^T & G P_a^T R_{2a}^{-1} R_{2a}^{-1} P_a G^T \end{bmatrix} > 0.$$

Then (A_c, B_c, C_c) given by

$$A_c = \Gamma(A - Q_a V_2^{-1} C - B R_{2a}^{-1} P_a + D P) G^T, \quad (5.2.5)$$

$$B_c = \Gamma Q_a V_2^{-1}, \quad (5.2.6)$$

$$C_c = -R_{2a}^{-1} P_a G^T \quad (5.2.7)$$

solves the dynamic robust stabilization problem.

Proof: In Theorem 4.1.1 define

$$\Phi(\bar{P}) = \bar{R} + \sum_{i=1}^p \delta_i (\bar{E}_i^T \bar{E}_i + \bar{P} \bar{D}_i \bar{D}_i^T \bar{P}), \quad \bar{M}_i = \bar{E}_i, \quad \bar{N}_i = \bar{D}_i^T \bar{P},$$

$$\bar{D}_i = \begin{bmatrix} D_i & 0 \\ 0 & 0 \end{bmatrix}, \quad \bar{E}_i = \begin{bmatrix} E_i & F_i C_c \\ 0 & 0 \end{bmatrix}$$

and note that (5.2.2) and (5.2.4) correspond to (4.1.8). Also, note that (5.2.1) and (5.2.3) correspond to

$$0 = (\tilde{A} + \sum_{i=1}^p \delta_i \tilde{D}_i \tilde{D}_i^T \tilde{P}) \tilde{Q} + \tilde{Q} (\tilde{A} + \sum_{i=1}^p \delta_i \tilde{D}_i \tilde{D}_i^T \tilde{P})^T + \tilde{V}. \quad (5.2.8)$$

Remark 5.2.1: Theorem 5.2.1 was discovered by optimizing $J = \text{tr } \tilde{P} \tilde{V}$ subject to

$$0 = \tilde{A}^T \tilde{P} + \tilde{P} \tilde{A} + \tilde{R} + \sum_{i=1}^p \delta_i [\tilde{E}_i^T \tilde{E}_i + \tilde{P} \tilde{D}_i \tilde{D}_i^T \tilde{P}] \quad (5.2.9)$$

and using the techniques of [40], [41]. As shown in [54], this approach also yields robust performance bounds.

ACKNOWLEDGMENT

The author wishes to thank Prof. C. V. Hollot for numerous helpful discussions and for carefully reviewing the original manuscript of the paper.

REFERENCES

- [1] P. K. Wong and M. Athans, "Closed-loop structural stability for linear-quadratic optimal systems," *IEEE Trans. Automat. Contr.*, vol. AC-22, pp. 94-99, 1977.
- [2] M. G. Safonov and M. Athans, "Gain and phase margin for multiloop LQG regulators," *IEEE Trans. Automat. Contr.*, vol. AC-22, pp. 173-179, 1977.
- [3] J. C. Doyle, "Guaranteed margins for LQG regulators," *IEEE Trans. Automat. Contr.*, vol. AC-23, pp. 756-757, 1978.
- [4] J. C. Doyle and G. Stein, "Robustness and observers," *IEEE Trans. Automat. Contr.*, vol. AC-24, pp. 607-611, 1979.
- [5] —, "Multivariable feedback design: Concepts for a classical/modern synthesis," *IEEE Trans. Automat. Contr.*, vol. AC-26, pp. 4-16, 1981.
- [6] G. Stein and M. Athans, "The LQG/LTR procedure for multivariable feedback control design," *IEEE Trans. Automat. Contr.*, vol. AC-32, pp. 105-114, 1987.
- [7] E. Soroka and U. Shaked, "On the robustness of LQ regulators," *IEEE Trans. Automat. Contr.*, vol. AC-29, pp. 664-665, 1984.
- [8] U. Shaked and E. Soroka, "On the stability robustness of the continuous-time LQG optimal control," *IEEE Trans. Automat. Contr.*, vol. AC-30, pp. 1039-1043, 1985.
- [9] M. Tahk and J. L. Speyer, "Modeling of parameter variations and asymptotic LQG synthesis," in *Proc. 25th IEEE Conf. Decision Contr.*, Athens, Greece, Dec. 1986, pp. 1459-1465.
- [10] S. S. L. Chang and T. K. C. Peng, "Adaptive guaranteed cost control of systems with uncertain parameters," *IEEE Trans. Automat. Contr.*, vol. AC-17, pp. 474-483, 1972.
- [11] R. V. Patel, M. Toda, and S. Sridhar, "Robustness of linear quadratic state feedback designs in the presence of system uncertainty," *IEEE Trans. Automat. Contr.*, vol. AC-22, pp. 945-949, 1977.
- [12] A. M. Meilakhs, "Design of stable control systems subject to parametric perturbation," *Automat. Remote Contr.*, pp. 1409-1418, 1979.
- [13] A. Vinkler and L. J. Wood, "Multistep guaranteed cost control of linear systems with uncertain parameters," *J. Guid. Contr.*, vol. 2, pp. 449-456, 1979.
- [14] M. Eslami and D. L. Russell, "On stability with large parameter variations: Stemming from the direct method of Lyapunov," *IEEE Trans. Automat. Contr.*, vol. AC-25, pp. 1231-1234, 1980.
- [15] M. Corless and G. Leitmann, "Continuous state feedback guaranteeing uniform ultimate boundedness for uncertain dynamical systems," *IEEE Trans. Automat. Contr.*, vol. AC-26, pp. 1139-1144, 1981.
- [16] J. S. Thorp and B. R. Barmish, "On guaranteed stability of uncertain linear systems via linear control," *J. Optimiz. Theory Appl.*, vol. 35, pp. 559-579, 1981.
- [17] E. Noldus, "Design of robust state feedback laws," *Int. J. Contr.*, vol. 35, pp. 935-944, 1982.
- [18] B. R. Barmish, I. R. Petersen, and A. Feuer, "Linear ultimate boundedness control of uncertain dynamic systems," *Automatica*, vol. 19, pp. 523-532, 1983.
- [19] B. R. Barmish, M. Corless, and G. Leitmann, "A new class of stabilizing controllers for uncertain dynamical systems," *SIAM J. Contr. Optimiz.*, vol. 21, pp. 246-255, 1983.
- [20] I. R. Petersen and B. R. Barmish, "Control effort considerations in the stabilization of uncertain dynamical systems," in *Proc. Amer. Contr. Conf.*, San Diego, CA, June 1984, pp. 490-495.
- [21] B. R. Barmish, "Necessary and sufficient conditions for quadratic stabilizability of an uncertain system," *J. Optimiz. Theory Appl.*, vol. 46, pp. 399-408, 1985.
- [22] R. K. Yedavalli, S. S. Banda, and D. B. Ridgely, "Time-domain stability robustness measures for linear regulators," *J. Guidance Contr. Dynam.*, vol. 8, pp. 520-524, 1985.
- [23] I. R. Petersen, "A Riccati equation approach to the design of stabilizing controllers and observers for a class of uncertain linear systems," *IEEE Trans. Automat. Contr.*, vol. AC-30, pp. 904-907, 1985.
- [24] A. R. Galimidi and B. R. Barmish, "The constrained Lyapunov problem and its application to robust output feedback stabilization," *IEEE Trans. Automat. Contr.*, vol. AC-31, pp. 410-419, 1986.
- [25] I. R. Petersen and C. V. Hollot, "A Riccati equation approach to the stabilization of uncertain systems," *Automatica*, vol. 22, pp. 397-411, 1986.
- [26] W. M. Wonham, "Optimal stationary control of a linear system with state-dependent noise," *SIAM J. Contr.*, vol. 5, pp. 486-500, 1967.
- [27] —, "On a matrix Riccati equation of stochastic control," *SIAM J. Contr.*, vol. 6, pp. 681-697, 1968.
- [28] D. Kleinman, "Optimal stationary control of linear systems with control-dependent noise," *IEEE Trans. Automat. Contr.*, vol. AC-14, pp. 673-677, 1969.
- [29] P. McLane, "Optimal stochastic control of linear systems with state and control-dependent disturbances," *IEEE Trans. Automat. Contr.*, vol. AC-16, pp. 793-798, 1971.
- [30] P. J. McLane, "Linear optimal stochastic control using instantaneous output feedback," *Int. J. Contr.*, vol. 13, pp. 383-396, 1971.
- [31] U. Haussmann, "Optimal stationary control with state and control dependent noise," *SIAM J. Contr.*, vol. 9, pp. 184-198, 1971.
- [32] C. W. Merriam, III, *Automated Design of Control Systems*. New York: Gordon and Breach, 1974.
- [33] J. Bismut, "Linear-quadratic optimal stochastic control with random coefficients," *SIAM J. Contr.*, vol. 14, pp. 419-444, 1976.
- [34] G. N. Mil'stein, "Design of stabilizing controller with incomplete state data for linear stochastic system with multiplicative noise," *Automat. Remote Contr.*, vol. 43, pp. 653-659, 1982.
- [35] Y. A. Phillis, "Controller design of systems with multiplicative noise," *IEEE Trans. Automat. Contr.*, vol. AC-30, pp. 1017-1019, 1985.
- [36] D. S. Bernstein and D. C. Hyland, "Optimal projection/maximum entropy stochastic modelling and reduced-order design synthesis," in *Proc. IFAC Workshop on Model Error Concepts and Compensation*, Boston, MA, June 1985, pp. 47-54.
- [37] —, "The optimal projection/maximum entropy approach to designing low-order, robust controllers for flexible structures," in *Proc. 24th IEEE Conf. Decision Contr.*, Fort Lauderdale, FL, Dec. 1985, pp. 745-752.
- [38] D. S. Bernstein, L. D. Davis, S. W. Greeley, and D. C. Hyland, "Numerical solution of the optimal projection/maximum entropy design equations for low-order, robust controller design," in *Proc. 24th IEEE Conf. Decision Contr.*, Fort Lauderdale, FL, Dec. 1985, pp. 1795-1798.
- [39] D. S. Bernstein and S. W. Greeley, "Robust controller synthesis using the maximum entropy design equations," *IEEE Trans. Automat. Contr.*, vol. AC-31, pp. 362-364, 1986.
- [40] D. S. Bernstein and D. C. Hyland, "The optimal projection equations for reduced-order modelling, estimation and control of linear systems with multiplicative white noise," *J. Optimiz. Theory Appl.*, 1988.
- [41] D. C. Hyland and D. S. Bernstein, "The optimal projection equations for fixed-order dynamic compensation," *IEEE Trans. Automat. Contr.*, vol. AC-29, pp. 1034-1037, 1984.
- [42] P. J. McLane, "Asymptotic stability of linear autonomous systems with state-dependent noise," *IEEE Trans. Automat. Contr.*, vol. AC-14, pp. 754-755, 1969.
- [43] J. L. Willems, "Mean square stability criteria for stochastic feedback systems," *Int. J. Syst. Sci.*, vol. 4, pp. 545-564, 1973.
- [44] U. G. Haussmann, "Stability of linear systems with control dependent noise," *SIAM J. Contr.*, vol. 11, pp. 382-394, 1973.
- [45] U. Haussmann, "On the existence of moments of stationary linear systems with multiplicative noise," *SIAM J. Contr.*, vol. 12, pp. 99-105, 1974.
- [46] J. L. Willems and J. C. Willems, "Feedback stabilizability for stochastic systems with state and control dependent noise," *Automatica*, vol. 12, pp. 277-283, 1976.
- [47] J. L. Willems and J. C. Willems, "Robust stabilization of uncertain systems," *SIAM J. Contr. Optimiz.*, vol. 21, pp. 352-374, 1983.
- [48] L. Arnold, H. Crauel, and V. Wihstutz, "Stabilization of linear systems by noise," *SIAM J. Contr. Optimiz.*, vol. 21, pp. 451-461, 1983.
- [49] B. D. O. Anderson and J. B. Moore, *Linear Optimal Control*. Englewood Cliffs, NJ: Prentice-Hall, 1970.
- [50] M. Athans, R. T. Ku, and S. B. Gershwin, "The uncertainty threshold principle: Some fundamental limitations of optimal decision making under dynamic uncertainty," *IEEE Trans. Automat. Contr.*, vol. AC-22, pp. 491-495, 1977.
- [51] R. T. Ku and M. Athans, "Further results on the uncertainty threshold

- principle," *IEEE Trans. Automat. Contr.*, vol. AC-22, pp. 866-868, 1977.
- [52] D. S. Bernstein and D. C. Hyland, "Optimal projection for uncertain systems (OPUS): A unified theory of reduced-order, robust control design," in *Large Space Structures: Dynamics and Control*, S. N. Atluri and A. K. Amos, Eds. New York: Springer-Verlag, 1987.
- [53] D. C. Hyland and D. S. Bernstein, "The majorant Lyapunov equation: A nonnegative matrix equation for guaranteed robust stability and performance of large scale systems," *Amer. Contr. Conf.*, Minneapolis, MN, June 1987, pp. 910-917; also in *IEEE Trans. Automat. Contr.*, vol. AC-32, pp. 1005-1013, Nov. 1987.
- [54] D. S. Bernstein and W. M. Haddad, "The optimal projection equations with Petersen-Holot bounds: Robust controller synthesis with guaranteed structured stability radius," *IEEE Conf. Decision Contr.*, Los Angeles, CA, Dec. 1987.
- [55] W. M. Haddad, "Robust optimal projection control-system synthesis," Ph.D. dissertation, Dep. Mechanical Eng., Florida Inst. Technol., Melbourne, FL, Mar. 1987.
- [56] J. W. Brewer, "Kronecker products and matrix calculus in system theory," *IEEE Trans. Circuits Syst.*, vol. CAS-25, pp. 772-781, 1978.
- [57] L. Arnold, *Stochastic Differential Equations: Theory and Applications*. New York: Wiley, 1974.
- [58] J. Medanic, "On stabilization and optimization by output feedback," in *Proc. Twelfth Asilomar Conf. Circ., Syst. Comp.*, 1978, pp. 412-416.
- [59] S. Renjen and D. P. Looze, "Synthesis of decentralized output/state regulators," in *Proc. Amer. Contr. Conf.*, Arlington, VA, June 1982, pp. 758-762.
- [60] D. S. Bernstein, "The optimal projection equations for static and dynamic output feedback: The singular case," this issue, pp. 1139-1143.
- [61] C. R. Rao and S. K. Mitra, *Generalized Inverse of Matrices and Its Applications*. New York: Wiley, 1971.
- [62] W. M. Wonham, *Linear Multivariable Control: A Geometric Approach*. New York: Springer-Verlag, 1979.

Dennis S. Bernstein (M'82), for a photograph and biography, see p. 1013 of the November 1987 issue of this TRANSACTIONS.

APPENDIX G: Robust Analysis

- 115. D. S. Bernstein and W. M. Haddad, "Robust Stability and Performance Analysis for Linear Dynamic Systems," IEEE Trans. Autom. Contr., Vol. 34, 1989, to appear.
- 123. D. S. Bernstein and W. M. Haddad, "Robust Stability and Performance Analysis for State Space Systems via Quadratic Lyapunov Bounds," submitted to SIAM J. Matrix Anal. Appl.
- 75. D. C. Hyland and D. S. Bernstein, "The Majorant Lyapunov Equation: A Nonnegative Matrix Equation for Guaranteed Robust Stability and Performance of Large Scale Systems," IEEE Trans. Autom. Contr., Vol. AC-32, pp. 1005-1013, 1987.

October 1987
Revised May 1988

**Robust Stability and Performance Analysis
for
Linear Dynamic Systems**

by

Dennis S. Bernstein
Harris Corporation
Government Aerospace Systems Division
MS 22/4848
Melbourne, FL 32902

Wassim M. Haddad
Department of Mechanical and
Aerospace Engineering
Florida Institute of Technology
Melbourne, FL 32901

Abstract

In a recent paper Zhou and Khargonekar obtained sufficient conditions for robust stability over specified sets of matrix perturbations. In the present paper these results are extended to include, in addition, performance bounds. Here performance is defined as the worst-case expected value of a quadratic functional involving the state variables when the system is subjected to white noise disturbances. The results are illustrated by considering the gain margin of both an LQG controller and a robustified design obtained by Bernstein and Greeley for Doyle's example.

1. Introduction

It is well known that unavoidable discrepancies between mathematical models and real-world systems can result in the degradation of control-system performance. Ideally, feedback control systems should be designed to be *robust* with respect to uncertainties in the plant characteristics. Thus robustness *analysis* must play a key role in control-system design. That is, given an existing or proposed control system, determine the performance degradation due to variations in the plant. The most fundamental concern in this regard is clearly that of stability. For linear state space systems with which the present paper is concerned, this problem has received increasing attention over the past several years (see, e.g., [1-12]).

One of the principal techniques used to assess robust stability is based upon quadratic Lyapunov functions (see [1-4,10]). Quadratic Lyapunov functions have also been used extensively for robust control-system *synthesis*; see [13] for relevant references. The problem of robust synthesis is, however, beyond the scope of the present paper.

In addition to assessing robust stability, it is often desirable to quantify performance by considering the degradation of a cost functional as the plant parameters deviate from their nominal values. Although any robustly stable system over a compact set of parameters possesses a worst-case performance, it is desirable in practice to actually determine a bound for the worst-case performance. The concern for *both* robust stability and performance goes back to the early work of Michael and Merriam [14], while more recent references include the work of Chang and Peng [15], Noldus [16], and Petersen [17]. The results of [15-17] can be shown to depend upon a modified Lyapunov equation of the form

$$0 = AQ + QA^T + \hat{\Omega}(Q) + V, \quad (1.1)$$

where the operator $\hat{\Omega}(Q)$ is chosen to bound terms of the form $\Delta AQ + Q\Delta A^T$, where ΔA is an uncertain perturbation of the dynamics matrix A . Since robust performance per se was not discussed in [16,17], the work most closely related to the present paper is that of Chang and Peng [15]. They essentially show that consideration of (1.1) leads to a bound on worst-case performance. Although the development in [15] was carried out for full-state feedback, specialization of their approach to robust performance analysis is straightforward. A systematic, in-depth treatment of robust performance analysis involving the approach of [15] as well as other bounds is given in [18].

The starting point for the present paper is the recent paper by Zhou and Khargonekar [10]. By analyzing the Lyapunov equation they obtain a series of stability robustness tests which improve

significantly upon earlier work [2-4]. In the present paper we extend the results of [10] to obtain, in addition, a bound on worst-case performance. As in (1.1) we consider a Lyapunov equation of the form

$$0 = AQ + QA^T + \Omega + V, \quad (1.2)$$

where Ω bounds uncertainty terms of the form $\Delta AQ + Q\Delta A^T$. The principal difference between (1.1) and (1.2) is that Ω in (1.2) is a constant matrix independent of the solution Q . The case considered in [15] in which Ω is a function of Q is discussed in [18].

The cost functional used in the present paper to quantify robust performance is the trace of the output covariance of a system subjected to white noise disturbances. This measure of performance is identical in form to the standard performance criterion of LQG theory. Since we also obtain a bound for the state covariance matrix, our results yield bounds on the variances (mean square response levels) of system states. Although the results of [15] were obtained within a deterministic setting, it is easy to see that the performance criterion of [15] is also of this form.

The contents of the paper are as follows. After introducing notation at the end of this section we consider the robust stability and performance problems in Section 2. In Section 3 we present the main result (Theorem 3.1) which provides sufficient conditions for robust stability over a set of parameter variations along with a performance bound. In Section 4 we present a dual result (Theorem 4.1) in terms of the dual matrix P . This result serves two purposes. First, it clarifies connections with the previous literature where results are presented in terms of the quadratic Lyapunov function $V(x) = x^T Px$. And, second, we show that the dual performance bound may be much better than the primal bound (and vice versa) for particular problems. The results of Theorems 3.1 and 4.1 are given in terms of a robustness set \mathcal{U} which is a subset of a maximal set $\tilde{\mathcal{U}}$. Since $\tilde{\mathcal{U}}$ is defined implicitly, we provide explicit characterizations of subsets \mathcal{U} in Section 5. Here we restate the principal results of [2-4,10] which, in our context, correspond to particular characterizations of subsets of $\tilde{\mathcal{U}}$. We also introduce an additional subset of $\tilde{\mathcal{U}}$ which provides a new robust stability result. Finally, in Section 6 we consider a pair of illustrative examples. The first example, which was previously considered in [10], involves two uncertain parameters. It is shown that the new guaranteed stability region is considerably larger for certain parameter values than the regions given in [10] (see Figure 1). Furthermore, we obtain a robust performance bound, a result which has no counterpart in [10]. The second example involves controllers for a second-order open-loop unstable plant originally considered in [19] to demonstrate the lack of a guaranteed stability margin for LQG controllers. We apply Theorems 3.1 and 4.1 to analyze both the LQG design and

a robustified design obtained in [20]. We show that the new robust stability test is effective in the sense that the *guaranteed* gain margin for the robustified controller is a factor of 5 larger than the *actual* gain margin of the LQG design.

Notation

Note: All matrices have real entries

| | |
|---|---|
| $\mathbb{R}, \mathbb{R}^{r \times s}, \mathbb{R}^r, \mathbb{E}$ | real numbers, $r \times s$ real matrices, $\mathbb{R}^{r \times 1}$, expectation |
| I_r | $r \times r$ identity matrix |
| asymptotically stable matrix | matrix with eigenvalues in open left half plane |
| $\mathbb{S}^r, \mathbb{N}^r, \mathbb{P}^r$ | $r \times r$ symmetric, nonnegative-definite, positive-definite matrices |
| $Z_1 \geq Z_2, Z_1 > Z_2$ | $Z_1 - Z_2 \in \mathbb{N}^r, Z_1 - Z_2 \in \mathbb{P}^r, Z_1, Z_2 \in \mathbb{S}^r$ |
| $\text{tr } Z, Z^T, \text{co}$ | trace of Z , transpose of Z , convex hull |
| $\lambda_{\min}(Z), \lambda_{\max}(Z)$ | smallest and largest eigenvalues of $Z \in \mathbb{S}^r$ |
| $\ Z\ _s$ | spectral norm |
| $Z_{(i,j)}$ | (i,j) element of matrix Z |
| $Z \geq 0$ | $Z_{(i,j)} \geq 0, i, j = 1, \dots, r, Z \in \mathbb{R}^{r \times r}$ |
| $Z >> 0$ | $Z_{(i,j)} > 0, i, j = 1, \dots, r, Z \in \mathbb{R}^{r \times r}$ |
| $ Z _m$ | $\{ Z_{(i,j)} \}_{i,j=1}^r, Z \in \mathbb{R}^{r \times r}$ (matrix modulus) |

2. Robust Stability and Performance Problems

Let $\mathcal{U} \subset \mathbb{R}^{n \times n}$ denote a set of perturbations ΔA of the nominal dynamics matrix A . Throughout the paper it is assumed that A is asymptotically stable. We begin by considering the question of whether or not $A + \Delta A$ is asymptotically stable for all $\Delta A \in \mathcal{U}$.

Robust Stability Problem. Determine whether the linear system

$$\dot{x}(t) = (A + \Delta A)x(t), \quad t \in [0, \infty), \quad (2.1)$$

is asymptotically stable for all $\Delta A \in \mathcal{U}$.

The problem of robust performance involves a quadratic form $x^T(t)Rx(t)$, where $R \in \mathbb{N}^n$, when the system is subjected to a white noise disturbance $w(t)$ with nonnegative-definite intensity V . The matrix R can be viewed as a means for selecting output variables of interest while the matrix V can be used to specify disturbance levels.

Robust Performance Problem. For the disturbed linear system

$$\dot{x}(t) = (A + \Delta A)x(t) + w(t), \quad t \in [0, \infty), \quad (2.2)$$

determine a performance bound β satisfying

$$J(\mathcal{U}) \triangleq \sup_{\Delta A \in \mathcal{U}} \limsup_{t \rightarrow \infty} \mathbb{E}[x^T(t) R x(t)] \leq \beta. \quad (2.3)$$

The system (2.2) may, for example, denote a control system in closed-loop configuration subjected to external white noise disturbances (see Section 6). Such specializations are not required for this development, however. Note that $J(\mathcal{U})$ represents the worst case (over \mathcal{U}) of the average (over the white noise statistics) of quadratically weighted steady-state deviations of the state from the origin. Thus β represents an upper bound on selected output variances.

Of course, since R and V are only assumed to be nonnegative definite, there may be cases in which a finite performance bound β satisfying (2.3) exists while (2.1) is not asymptotically stable over \mathcal{U} . In practice, however, robust performance is mainly of interest when (2.1) is robustly stable. In this case the performance $J(\mathcal{U})$ is given in terms of the steady-state second moment of the state. The following result from linear system theory will be useful.

Lemma 2.1. Suppose (2.1) is asymptotically stable for all $\Delta A \in \mathcal{U}$. Then

$$J(\mathcal{U}) = \sup_{\Delta A \in \mathcal{U}} \text{tr } Q_{\Delta A} R, \quad (2.4)$$

where $n \times n$ $Q_{\Delta A} \triangleq \lim_{t \rightarrow \infty} \mathbb{E}[x(t)x^T(t)]$ is the unique, nonnegative-definite solution to

$$0 = (A + \Delta A)Q_{\Delta A} + Q_{\Delta A}(A + \Delta A)^T + V. \quad (2.5)$$

In the present paper our approach is to obtain sufficient conditions for robust stability as a *consequence* of sufficient conditions for robust performance. Such conditions are developed in the following sections.

3. Sufficient Conditions for Robust Stability and Performance

The key step in obtaining robust stability and performance is to replace the uncertain terms in the Lyapunov equation (2.5) by a bounding matrix Ω . The nonnegative-definite solution Q of this bounding Lyapunov equation is then guaranteed to be an upper bound for $Q_{\Delta A}$. The uncertainty

set \mathcal{U} over which robustness is guaranteed then depends upon Q . The following easily proved result is fundamental and forms the basis for all later developments. The hypotheses of this result are of a general nature and are not intended to be directly verifiable. Suitably verifiable specializations of the hypotheses are discussed in Section 5.

Theorem 3.1. Let $\Omega \in \mathbb{R}^n$, let $Q \in \mathbb{R}^n$ be the unique solution to

$$0 = AQ + QA^T + \Omega + V, \quad (3.1)$$

and let \mathcal{U} be a subset of

$$\tilde{\mathcal{U}} \triangleq \{\Delta A \in \mathbb{R}^{n \times n} : \Delta A Q + Q \Delta A^T \leq \Omega\}. \quad (3.2)$$

Then

$$(A + \Delta A, [V + \Omega - (\Delta A Q + Q \Delta A^T)]^{\frac{1}{2}}) \text{ is stabilizable, } \Delta A \in \mathcal{U}, \quad (3.3)$$

if and only if

$$A + \Delta A \text{ is asymptotically stable, } \Delta A \in \mathcal{U}. \quad (3.4)$$

In this case,

$$Q_{\Delta A} \leq Q, \quad \Delta A \in \mathcal{U}, \quad (3.5)$$

where $Q_{\Delta A} \in \mathbb{R}^n$ is given by (2.5), and

$$J(\mathcal{U}) \leq \text{tr } QR. \quad (3.6)$$

If, in addition, there exists $\Delta A \in \tilde{\mathcal{U}}$ such that $(A + \Delta A, [V + \Omega - (\Delta A Q + Q \Delta A^T)]^{\frac{1}{2}})$ is controllable, then Q is positive definite.

Proof. This result is a minor variation of Theorem 3.1 of [21] and hence the proof is omitted. \square

To apply Theorem 3.1, one first chooses a nonnegative-definite matrix Ω and then solves (3.1) for Q . Next, as shown in Section 4, one examines $\tilde{\mathcal{U}}$ to determine subsets \mathcal{U} of perturbations ΔA over which robustness is guaranteed. Note that if \mathcal{U}_1 and \mathcal{U}_2 are subsets of $\tilde{\mathcal{U}}$ then so is the convex hull of their union. (To see this note that $\tilde{\mathcal{U}}$ is convex.) The set $\tilde{\mathcal{U}}$ is the largest set over which robustness can be guaranteed by Theorem 3.1 for the particular choice of Ω . One may also select several matrices Ω and determine subsets of each resulting $\tilde{\mathcal{U}}$ as a constructive approach to

determining larger robustness sets. In the next section we examine subsets \mathcal{U} of $\tilde{\mathcal{U}}$ of specified structure. Before doing so, we have the following observations.

In applying Theorem 3.1 it may be convenient to replace condition (3.3) with stronger conditions which are easier to verify in practice. The following result is immediate.

Proposition 3.1. Consider the conditions

$$V > 0, \quad (3.7)$$

$$(A + \Delta A, V^{\frac{1}{2}}) \text{ is stabilizable, } \Delta A \in \mathcal{U}, \quad (3.8)$$

$$\Delta A Q + Q \Delta A^T < \Omega, \quad \Delta A \in \mathcal{U}, \quad (3.9)$$

$$\Delta A Q + Q \Delta A^T < \Omega + V, \quad \Delta A \in \mathcal{U}. \quad (3.10)$$

Then (3.7) \Rightarrow (3.8) \Rightarrow (3.3), (3.7) \Rightarrow (3.10) \Rightarrow (3.3), and (3.9) \Rightarrow (3.10) \Rightarrow (3.3).

If only robust stability is of interest, then the noise intensity V need not have physical significance. In this case one may either set $V = \epsilon I_n$, where $\epsilon > 0$ is small to satisfy (3.7), or set $V = 0$ and confine \mathcal{U} to perturbations ΔA for which (3.9) holds. This is the case in [3,4,10] where $V = 0, \Omega = 2I_n$, and the parametric robustness sets are characterized by strict inequality.

Remark 3.1. Since A is asymptotically stable, Q is given by

$$Q = \int_0^\infty e^{At}(\Omega + V)e^{A^T t} dt = \int_0^\infty e^{At}\Omega e^{A^T t} dt + Q_0, \quad (3.11)$$

where $Q_0 \in \mathbb{R}^n$ is given by

$$0 = A Q_0 + Q_0 A^T + V. \quad (3.12)$$

Note that $Q_0 \leq Q$ and that the nominal performance is given by $\text{tr } Q_0 R$.

Remark 3.2. Using (3.11) it is also useful to note that the bound for $J(\mathcal{U})$ given by (3.6) can be written as

$$\text{tr } QR = \text{tr } \int_0^\infty e^{At}(\Omega + V)e^{A^T t} dt R = \text{tr } P_0(\Omega + V), \quad (3.13)$$

where $P_0 \in \mathbb{R}^n$ is given by

$$0 = A^T P_0 + P_0 A + R. \quad (3.14)$$

The bound $\text{tr } P_0(\Omega + V)$ can be viewed as a dual formulation of the bound $\text{tr } QR$ since the roles of A and A^T are reversed. Dual bounds are developed in the following section. Note that $\text{tr } Q_0 R = \text{tr } P_0 V$.

4. Dual Sufficient Conditions for Robust Stability and Performance

As noted in Remark 3.2, the performance bound $\text{tr } QR$ given by (3.6) can be expressed equivalently in terms of a dual variable P_0 for which the roles of A and A^T are reversed. Using a similar technique, additional conditions for robust stability and performance can be obtained by developing a dual version of Theorem 3.1. A prime motivation for developing such dual bounds is to draw direct connections with previous results in the literature relating to robust stability. Traditionally, the use of the quadratic Lyapunov function $V(x) = x^T P x$ for robust stability leads naturally to the dual formulation. In addition, the dual bounds may, for certain problems, be much sharper than the bounds introduced in the previous section. This point is illustrated at the end of this section by examining an extreme case and in Section 6 by means of numerical examples. We note, in addition, that robust performance bounds are more difficult to motivate within the dual formulation without first developing the primal results. The following result is immediate.

Lemma 4.1. Suppose (2.1) is asymptotically stable for all $\Delta A \in \mathcal{U}$. Then

$$J(\mathcal{U}) = \sup_{\Delta A \in \mathcal{U}} \text{tr } P_{\Delta A} V, \quad (4.1)$$

where $n \times n$ $P_{\Delta A}$ is the unique, nonnegative-definite solution to

$$0 = (A + \Delta A)^T P_{\Delta A} + P_{\Delta A} (A + \Delta A) + R. \quad (4.2)$$

The dual of Theorem 3.1 can now be stated.

Theorem 4.1. Let $\Lambda \in \mathbb{R}^n$, let $P \in \mathbb{R}^n$ be the unique solution to

$$0 = A^T P + P A + \Lambda + R, \quad (4.3)$$

and let \mathcal{U} be a subset of

$$\tilde{\mathcal{U}}' \triangleq \{\Delta A \in \mathbb{R}^{n \times n} : \Delta A^T P + P \Delta A \leq \Lambda\}. \quad (4.4)$$

Then

$$\left([R + \Lambda - (\Delta A^T P + P \Delta A)]^{\frac{1}{2}}, A + \Delta A \right) \text{ is detectable, } \Delta A \in \mathcal{U}, \quad (4.5)$$

if and only if

$$A + \Delta A \text{ is asymptotically stable, } \Delta A \in \mathcal{U}. \quad (4.6)$$

In this case,

$$P_{\Delta A} \leq P, \quad \Delta A \in \mathcal{U}, \quad (4.7)$$

where $P_{\Delta A} \in \mathbb{R}^n$ is given by (4.2), and

$$J(\mathcal{U}) \leq \text{tr } PV. \quad (4.8)$$

If, in addition, there exists $\Delta A \in \mathcal{U}$ such that $([R + \Lambda - (\Delta A^T P + P \Delta A)]^{\frac{1}{2}}, A + \Delta A)$ is observable, then P is positive definite.

The usefulness of Theorem 4.1 resides in the fact that it provides stability and performance bounds which are generally different from those given by Theorem 3.1. Hence, depending upon Ω and Λ either bound (3.6) or bound (4.8) may be better for a particular problem. To illustrate how dual bounds can improve estimates of robust performance, consider the case in which $V = 0$, i.e., plant disturbances are absent. In this case $Q_{\Delta A} = 0$ satisfies (2.5) and thus $J(\mathcal{U}) = 0$ so long as $A + \Delta A$ is stable for all $\Delta A \in \mathcal{U}$. The performance bound $\text{tr } QR$ given by (3.6) may, however, be arbitrarily large depending upon R since Q may be nonzero due to Ω . Hence this performance bound may be arbitrarily conservative. The dual bound (4.8), on the other hand, is zero in this case, which completely eliminates the conservatism.

5. Characterization of Subsets of $\tilde{\mathcal{U}}$ and $\tilde{\mathcal{U}}'$

To apply Theorems 3.1 and 4.1 it is necessary to explicitly characterize subsets \mathcal{U} of $\tilde{\mathcal{U}}$ and $\tilde{\mathcal{U}}'$ over which robustness is guaranteed. In this section we provide several such characterizations by collecting together and extending known results from the literature.

For the following result let $\Omega = \omega I_n$, where $\omega > 0$, let $W \in \mathbb{R}^{n \times n}$, $W \gg 0$, and let $A_1, \dots, A_p \in \mathbb{R}^{n \times n}$ be arbitrary. Furthermore, for $Q \in \mathbb{P}^n$ satisfying (3.1) define for $i = 1, \dots, p$:

$$\alpha_i \triangleq \lambda_{\min}(A_i Q + Q A_i^T), \quad \beta_i \triangleq \lambda_{\max}(A_i Q + Q A_i^T),$$

$$\begin{aligned} I_i &\triangleq (-\infty, \infty), & \alpha_i &= \beta_i = 0, \\ &\triangleq (-\infty, \omega/\beta_i), & \alpha_i &\geq 0, \beta_i > 0, \\ &\triangleq (\omega/\alpha_i, \infty), & \alpha_i &< 0, \beta_i \leq 0, \\ &\triangleq (\omega/\alpha_i, \omega/\beta_i), & \alpha_i &< 0 < \beta_i. \end{aligned}$$

Finally, let $e_i^{(p)}$ denote the i th column of the $p \times p$ identity matrix.

Proposition 5.1. Let $Q \in \mathbb{P}^n$ satisfy (3.1) with $\Omega = \omega I_n$, where $\omega > 0$. Then the following sets are subsets of \tilde{U} which also satisfy (3.9):

$$\mathcal{U}_1 \triangleq \{\Delta A \in \mathbb{R}^{n \times n} : \|\Delta A\|_* < \frac{\omega}{2} \|Q\|_*^{-1}\},$$

$$\mathcal{U}_2 \triangleq \{\Delta A \in \mathbb{R}^{n \times n} : |\Delta A|_m < \omega \|W|Q|_m + |Q|_m W^T\|_*^{-1} W\},$$

$$\mathcal{U}_3 \triangleq \{\Delta A \in \mathbb{R}^{n \times n} : \Delta A = \sum_{i=1}^p \sigma_i A_i, \quad (\sigma_1, \dots, \sigma_p)^T \in \mathcal{R}\},$$

where \mathcal{R} is one of the following regions in \mathbb{R}^p :

$$\mathcal{R}_1 \triangleq \{(\sigma_1, \dots, \sigma_p) : \sum_{i=1}^p |\sigma_i| \|A_i Q + Q A_i^T\|_* < \omega\},$$

$$\mathcal{R}_2 \triangleq \{(\sigma_1, \dots, \sigma_p) : \sum_{i=1}^p \sigma_i^2 < \omega^2 \left\| \sum_{i=1}^p (A_i Q + Q A_i^T)^2 \right\|_*^{-1}\},$$

$$\mathcal{R}_3 \triangleq \{(\sigma_1, \dots, \sigma_p) : |\sigma_i| < \omega \left\| \sum_{i=1}^p |A_i Q + Q A_i^T|_m \right\|_*^{-1}, \quad i = 1, \dots, p\},$$

$$\mathcal{R}_4 \triangleq \text{co}\{\sigma_i e_i^{(p)} : \sigma_i \in I_i, \quad i = 1, \dots, p\}.$$

For the dual case we set $\Lambda = \lambda I_n$, where $\lambda > 0$, and define the dual sets $\tilde{U}'_1, \tilde{U}'_2, \tilde{U}'_3, I'_1, \mathcal{R}'_1, \mathcal{R}'_2, \mathcal{R}'_3$, and \mathcal{R}'_4 in an analogous fashion.

Remark 5.1. The proof of Proposition 5.1 is omitted since the results are either known or are immediate. Specifically, \mathcal{U}'_1 can be found in [2] while \mathcal{U}'_2 appears in [3,4]. The sets $\mathcal{R}'_1, \mathcal{R}'_2$, and \mathcal{R}'_3 are given in [10]. The set \mathcal{R}'_4 has not appeared previously in the literature although the result is immediate. It is only necessary to diagonalize $A_i^T P + P A_i$ by means of an orthogonal transformation and compare diagonal elements to obtain I'_i . Taking the convex hull over the intervals I'_i thus yields \mathcal{R}'_4 . Of course, the required eigenproblem entails additional computation.

Remark 5.2. Although most of the dual of Proposition 5.1 has appeared previously, the primal result Proposition 5.1 has not been discussed in the literature. For robust stability this result can be obtained by considering the stability of A^T in place of A . As will be shown in Section 6, the primal and dual results lead in general to different robust stability regions and performance bounds. It should also be stressed that although most of the dual of Proposition 5.1 has appeared previously, the present paper extends its applicability to the problem of robust performance in addition to robust stability.

Remark 5.3. As mentioned previously, the convex hull of the union of any collection of subsets of \tilde{U} is also a subset of \tilde{U} since \tilde{U} is convex. This observation applies to \mathcal{U}_3 in the sense that if \mathcal{U}_3 is

a subset of \tilde{U} with regions $\mathcal{R} = \hat{\mathcal{R}}$ and $\mathcal{R} = \hat{\mathcal{R}}$ separately, then \mathcal{U}_3 is also a subset with \mathcal{R} equal to the convex hull of the union of $\hat{\mathcal{R}}$ and $\hat{\mathcal{R}}$. Note that these observations follow from the convexity of \tilde{U} and do not contradict the fact that the set of asymptotically stable matrices is not convex.

Remark 5.4. The requirement that Ω be of the form ωI_n is not a constraint in applying Proposition 5.1. Indeed, it is only required that Ω be positive definite. To see this let invertible $\phi \in \mathbb{R}^{n \times n}$ be such that $\phi \Omega \phi^T = I_n$. Then Proposition 5.1 can be applied with suitable transformations of $\Delta A, Q, W$, and A_i .

Remark 5.5. As in [2-4,10], the sets $\mathcal{U}_1, \mathcal{U}_2, \mathcal{R}_1, \mathcal{R}_2$, and \mathcal{R}_3 are defined in terms of strict inequalities. In this case $\mathcal{U}_1, \mathcal{U}_2$, and \mathcal{U}_3 consist of elements of \tilde{U} satisfying $\Delta A Q + Q \Delta A^T < \Omega$ so that (3.9) is satisfied. Thus, by Proposition 3.1, the stabilizability condition (3.3) is automatically satisfied without reference to V .

Remark 5.6. In the special case $p = 1$ it is clear that $\mathcal{R}_1 = \mathcal{R}_2$. Furthermore, in this case \mathcal{R}_3 is always a subset of \mathcal{R}_1 and \mathcal{R}_2 and hence leads to a more conservative stability region. The largest possible set of perturbations ΔA of the form $\sigma_i A_i$ contained in \tilde{U} is given by \mathcal{R}_4 .

Remark 5.7. It is shown in Remark 2.12 of [10] that \mathcal{U}_2 can be obtained as a consequence of \mathcal{U}_3 with $\mathcal{R} = \mathcal{R}_3$ and a suitable choice of A_i . Hence \mathcal{U}_2 need not actually be considered separately. Our assumption that $W \gg 0$ (and not $W \geq 0$) is for convenience only.

Remark 5.8. Note that all of the subsets of \tilde{U} given by Proposition 5.1 are symmetric except for \mathcal{U}_3 with $\mathcal{R} = \mathcal{R}_4$. When the actual stability region is highly asymmetric, it follows that a symmetric robust stability region is necessarily highly conservative. This observation is illustrated by an example in Section 6.

Remark 5.9. The regions given by $\mathcal{R}_1, \mathcal{R}_2$, and \mathcal{R}_3 correspond, respectively, to 1-norm, 2-norm, and ∞ -norm neighborhoods. These results can easily be extended to include more general regions. For example, in the definition of \mathcal{R}_2 replace σ_i by σ_i/a_i and $A_i^T Q + Q A_i^T$ by $a_i(A_i^T Q + Q A_i^T)$, where a_i is an arbitrary positive constant, $i = 1, \dots, p$. With this modification \mathcal{R}_2 corresponds to an elliptical robust stability region. Detailed investigation of such regions is beyond the scope of this paper.

Remark 5.10. When each interval I_i is finite, or when only a finite interval is of interest, \mathcal{R}_4 can be expressed as the convex hull of a finite number of points. Specifically, letting $I_i =$

$[a_i, b_i]$, $i = 1, \dots, p$, it follows that

$$\mathcal{R}_4 = \text{co}\{a_1 e_1^{(p)}, b_1 e_1^{(p)}, \dots, a_p e_p^{(p)}, b_p e_p^{(p)}\}.$$

This set is illustrated by means of an example in the next section.

6. Examples

As a first example we adopt Example 2 of [10]. This example, which involves two uncertain parameters, was used in [10] to illustrate the robust stability regions \mathcal{R}'_1 , \mathcal{R}'_2 , and \mathcal{R}'_3 . The problem was originally cast in the form of a static output feedback controller with uncertain gains. Here for convenience in discussing robust performance we reformulate the example to involve uncertainty in the control input matrix. Hence consider the control system

$$\dot{x}(t) = A_0 x(t) + B_0 u(t), \quad (6.1)$$

$$y(t) = C_0 x(t), \quad (6.2)$$

$$u(t) = K y(t), \quad (6.3)$$

where

$$A_0 = \begin{bmatrix} -1 & 0 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & -3 \end{bmatrix}, \quad B_0 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 1 \end{bmatrix}, \quad C_0 = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}, \quad K = \begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix},$$

and the uncertainty ΔB_0 in B_0 is given by

$$\Delta B_0 = \begin{bmatrix} -\sigma_1 & 0 \\ 0 & -\sigma_2 \\ -\sigma_1 & -\sigma_2 \end{bmatrix}.$$

The closed-loop dynamics matrix is then given by

$$A + \Delta A = \begin{bmatrix} -2 + \sigma_1 & 0 & -1 + \sigma_1 \\ 0 & -3 + \sigma_2 & 0 \\ -1 + \sigma_1 & -1 + \sigma_2 & -4 + \sigma_1 \end{bmatrix},$$

where $\Delta A = \sigma_1 A_1 + \sigma_2 A_2$ and A_1, A_2 have the evident definitions. It can easily be shown that the exact stability region is given by $\sigma_1 \in (-\infty, 1.75)$ and $\sigma_2 \in (-\infty, 3)$. Thus the nominal dynamics matrix corresponding to $\sigma_1 = \sigma_2 = 0$ lies in the upper right-hand corner of the exact stability region so that, as noted in Remark 5.8, a high degree of conservatism can be expected using symmetric robustness regions. To consider robust stability alone, set $V = R = 0$ and $\omega = \lambda = 2$. In this case

regions $\mathcal{R}'_1, \mathcal{R}'_2$, and \mathcal{R}'_3 , as computed in [10], are shown in Figure 1. Region \mathcal{R}'_4 for this problem is given (see Remark 5.10) by

$$\mathcal{R}'_4 = \text{co} \left\{ \begin{pmatrix} -29.6 \\ 0 \end{pmatrix}, \begin{pmatrix} 1.65 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ -20.5 \end{pmatrix}, \begin{pmatrix} 0 \\ 2.85 \end{pmatrix} \right\},$$

which accounts somewhat better for the asymmetry of the stability region. The regions $\mathcal{R}_1, \mathcal{R}_2$, and \mathcal{R}_3 were found to be smaller than the corresponding dual regions, while \mathcal{R}_4 is given by

$$\mathcal{R}_4 = \text{co} \left\{ \begin{pmatrix} -31.1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1.64 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ -10.4 \end{pmatrix}, \begin{pmatrix} 0 \\ 2.63 \end{pmatrix} \right\},$$

which yields slight improvement in σ_1 .

To evaluate robust performance replace (6.1) by

$$\dot{x}(t) = A_0 x(t) + B_0 u(t) + w(t), \quad (6.4)$$

and define

$$J = \lim_{t \rightarrow \infty} \mathbb{E} [x^T(t) R_1 x(t) + u^T(t) R_2 u(t)],$$

which corresponds to (2.3) with $R = R_1 + C_0^T K^T R_2 K C_0$. Hence setting $R_1 = I_3$ and $R_2 = I_2$ yields

$$R = \begin{bmatrix} 2 & 0 & 1 \\ 0 & 2 & 0 \\ 1 & 0 & 2 \end{bmatrix}.$$

We also set $V = I_3$ and $\omega = \lambda = 2$. The resulting stability region for these values of V and R is given by

$$\begin{aligned} \mathcal{R}'_1 &= \{(\sigma_1, \sigma_2) : |\sigma_1|/.70 + |\sigma_2|/1.46 < 1\}, \\ \mathcal{R}'_2 &= \{(\sigma_1, \sigma_2) : \sigma_1^2 + \sigma_2^2 < (.70)^2\}, \\ \mathcal{R}'_3 &= \{(\sigma_1, \sigma_2) : |\sigma_1| < .68, \quad i = 1, 2\}, \\ \mathcal{R}'_4 &= \text{co} \left\{ \begin{pmatrix} -20.5 \\ 0 \end{pmatrix}, \begin{pmatrix} .70 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ -13.7 \end{pmatrix}, \begin{pmatrix} 0 \\ 1.46 \end{pmatrix} \right\}. \end{aligned}$$

Over these combined regions the performance bound was computed to be $\text{tr } PV = 2.26$. The primal result produced the regions

$$\begin{aligned} \mathcal{R}_1 &= \{(\sigma_1, \sigma_2) : |\sigma_1|/1.09 + |\sigma_2|/1.75 < 1\}, \\ \mathcal{R}_2 &= \{(\sigma_1, \sigma_2) : \sigma_1^2 + \sigma_2^2 < (1.08)^2\}, \\ \mathcal{R}_3 &= \{(\sigma_1, \sigma_2) : |\sigma_1| < 1.0, \quad i = 1, 2\}, \\ \mathcal{R}_4 &= \text{co} \left\{ \begin{pmatrix} -20.8 \\ 0 \end{pmatrix}, \begin{pmatrix} 1.09 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ -6.93 \end{pmatrix}, \begin{pmatrix} 0 \\ 1.75 \end{pmatrix} \right\}. \end{aligned}$$

Over these regions the performance bound was computed to be $\text{tr } QR = 3.18$. Contour plots of actual performance for perturbed values of σ_1 and σ_2 are shown in Figure 2. Note that when determining robust performance Theorems 3.1 and 4.1 yield performance bounds over robust stability regions which are generally smaller than the robust stability regions determined with $R = 0$ and $V = 0$. This mechanism represents the natural tradeoff between stability and performance. In general, to determine the largest stability regions, V and R should be set to zero initially.

As a second example we consider the control system given in [19] to demonstrate the lack of guaranteed gain margin for LQG controllers. Hence consider

$$\dot{x}_0(t) = A_0 x_0(t) + B_0 u(t) + w_1(t), \quad (6.5)$$

$$y(t) = C_0 x_0(t) + w_2(t), \quad (6.6)$$

with controller

$$\dot{x}_c(t) = A_c x_c(t) + B_c y(t), \quad (6.7)$$

$$u(t) = C_c x_c(t), \quad (6.8)$$

and performance

$$J = \lim_{t \rightarrow \infty} \mathbb{E} [x_0^T(t) R_1 x_0(t) + u^T(t) R_2 u(t)]. \quad (6.9)$$

The data are

$$A_0 = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}, \quad B_0 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \quad C_0 = [1 \quad 0],$$

$$V_1 = R_1 = \rho \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}, \quad V_2 = R_2 = 1,$$

where V_1 and V_2 are the intensities of $w_1(t)$ and $w_2(t)$, respectively. Uncertainty ΔB_0 in B_0 is thus represented by $\sigma_1 B_1$, where $B_1 = [0 \ 1]^T$. Thus, the closed-loop system corresponds to

$$A = \begin{bmatrix} A_0 & B_0 C_c \\ B_c C_0 & A_c \end{bmatrix}, \quad A_1 = \begin{bmatrix} 0 & B_1 C_c \\ 0 & 0 \end{bmatrix},$$

$$R = \begin{bmatrix} R_1 & 0 \\ 0 & 0 \end{bmatrix}, \quad V = \begin{bmatrix} V_1 & 0 \\ 0 & B_c V_2 B_c^T \end{bmatrix},$$

where the zero in the (2,2) block of R denotes the fact that we are considering the robust performance bound for the state regulation cost only. Choosing $\rho = 60$, it follows that the LQG gains are given by

$$A_c = \begin{bmatrix} -9 & 1 \\ -20 & -9 \end{bmatrix}, \quad B_c = \begin{bmatrix} 10 \\ 10 \end{bmatrix}, \quad C_c = [-10 \quad -10].$$

For this controller the actual stability region corresponds to $\sigma_1 \in (-.07, .01)$ (see Figure 3). Applying the results of Section 5 with $V = R = 0$ (for robust stability only) and $\omega = \lambda = 2$, we obtain

$$\begin{aligned}\mathcal{R}_1 = \mathcal{R}_2 = \mathcal{R}_3 &= (-.000242, .000242), & \mathcal{R}_4 &= (-.000242, .000728), \\ \mathcal{R}'_1 = \mathcal{R}'_2 &= (-.0000247, .0000247), & \mathcal{R}'_3 &= (-.0000219, .0000219), \\ \mathcal{R}'_4 &= (-.0000247, .0000265).\end{aligned}$$

Note that although the primal results are better than the dual results by an order of magnitude, they are conservative by two orders of magnitude with respect to the actual gain margin. For robust performance we again set $\omega = \lambda = 2$ and, using R and V given above, we obtained the bound $\text{tr } QR = 7633$ over the stability region $\mathcal{R}_4 = (-.000192, .000613)$. The nominal performance was given by $\text{tr } Q_0 R = \text{tr } P_0 V = 4875$, while the dual performance bound was $\text{tr } PV = 10510$ over $\mathcal{R}'_4 = (-.0000222, .0000238)$.

Robustified controllers for the example of [19] were obtained in [20] using the approach discussed in [13]. As shown in Figure 3 (see also [20]), the closed-loop system with the controller

$$A_c = \begin{bmatrix} -10.69 & 1 \\ -32.97 & -5.295 \end{bmatrix}, \quad B_c = \begin{bmatrix} 11.69 \\ 26.67 \end{bmatrix}, \quad C_c = [-6.245 \quad -6.245],$$

is stable over the range $\sigma_1 \in (-.28, .21)$. Hence we wish to determine whether the robust stability tests are capable of detecting this increase in gain margin. Applying Theorems 3.1 and 4.1 with $\omega = \lambda = 2$ and $V = R = 0$ yields stability for σ_1 in the regions $\mathcal{R}_1 = \mathcal{R}_2 = \mathcal{R}_3 = (-.0115, .0115)$ and $\mathcal{R}_4 = (-.0115, .057)$. This guarantee of stability is two orders of magnitude greater than the guarantee for the LQG design but is still an order of magnitude conservative with respect to the actual stability region for this controller. Note, however, that for $\sigma_1 > 0$ the guaranteed gain margin for the robustified design given by \mathcal{R}_4 (i.e., .057) is greater than the *actual* gain margin of the LQG design (.01). Hence the robustness test given by the \mathcal{R}_4 was able to detect a factor of 5 stability augmentation provided by the robustified design compared to the LQG controller. Finally, the robust performance bound for this controller was computed to be $\text{tr } QR = 11185$ over the region $\mathcal{R}_4 = (-.00165, .00493)$, while the dual bound was found to be $\text{tr } PV = 11223$ over $\mathcal{R}'_4 = (-.000724, .00123)$. For this problem the nominal performance is $\text{tr } Q_0 R = \text{tr } P_0 V = 9997$.

Acknowledgment. We wish to thank J. Straehla for wordprocessing support, T. Rhodes for preparing Figure 1, A. Tellez for performing the numerical calculations, and A. Daubendiek for producing Figures 2 and 3.

References

1. M. Eslami and D. L. Russell, "On Stability with Large Parameter Variations Stemming from the Direct Method of Lyapunov," *IEEE Trans. Autom. Contr.*, Vol. AC-25, pp. 1231-1234, 1980.
2. R. V. Patel and M. Toda, "Quantitative Measures of Robustness for Multivariable Systems," in *Proc. Joint. Autom. Contr. Conf.*, paper TP8-A, San Francisco, CA, June 1980.
3. R. K. Yedavalli, "Improved Measures of Stability Robustness for Linear State Space Models," *IEEE Trans. Autom. Contr.*, Vol. AC-30, pp. 577-579, 1985.
4. R. K. Yedavalli, "Perturbation Bounds for Robust Stability in Linear State Space Models," *Int. J. Contr.*, Vol. 42, pp. 1507-1517, 1985.
5. C. Van Loan, "How Near is a Stable Matrix to an Unstable Matrix?," *Contemporary Mathematics*, Vol. 47, pp. 465-478, American Mathematical Society, 1985.
6. D. Hinrichsen and A. J. Pritchard, "Stability Radii of Linear Systems," *Sys. Contr. Lett.*, Vol. 7, pp. 1-10, 1986.
7. D. Hinrichsen and A. J. Pritchard, "Stability Radii for Structured Perturbations and the Algebraic Riccati Equations," *Sys. Contr. Lett.*, Vol. 8, pp. 105-113, 1987.
8. J. M. Martin and G. A. Hewer, "Smallest Destabilizing Perturbations for Linear Systems," *Int. J. Contr.*, Vol. 45, pp. 1495-1504, 1987.
9. R. M. Biernacki, H. Hwang, and S. P. Bhattacharyya, "Robust Stability with Structured Real Parameter Perturbations," *IEEE Trans. Autom. Contr.*, Vol. AC-32, pp. 495-506, 1987.
10. K. Zhou and P. P. Khargonekar, "Stability Robustness Bounds for Linear State-Space Models with Structured Uncertainty," *IEEE Trans. Autom. Contr.*, Vol. AC-32, pp. 621-623, 1987.
11. D. C. Hyland and D. S. Bernstein, "The Majorant Lyapunov Equation: A Nonnegative Matrix Equation for Guaranteed Robust Stability and Performance of Large Scale Systems," *IEEE Trans. Autom. Contr.*, Vol. AC-32, pp. 1005-1113, 1987.
12. D. D. Siljak, "Parameter Space Methods for Robust Control Design: A Survey," preprint.
13. D. S. Bernstein, "Robust Output-Feedback Stabilization: Deterministic and Stochastic Perspectives," *IEEE Trans. Autom. Contr.*, Vol. AC-32, pp. 1076-1084, 1987.
14. G. S. Michael and C. W. Merriam, "Stability of Parametrically Disturbed Linear Optimal Control Systems," *J. Math. Anal. Appl.*, Vol. 28, pp. 294-302, 1969.
15. S. S. L. Chang and T. K. C. Peng, "Adaptive Guaranteed Cost Control of Systems with Uncertain Parameters," *IEEE Trans. Autom. Contr.*, Vol. AC-17, pp. 474-483, 1972.
16. E. Noldus, "Design of Robust State Feedback Laws," *Int. J. Contr.*, Vol. 35, pp. 935-944, 1982.
17. I. R. Petersen, "A Riccati Equation Approach to the Design of Stabilizing Controllers and Observers for a Class of Uncertain Linear Systems," *IEEE Trans. Autom. Contr.*, Vol. AC-30, pp. 904-907, 1985.
18. D. S. Bernstein and W. M. Haddad, "Robust Stability and Performance Analysis of State Space Systems with Structured Uncertainty via Quadratic Lyapunov Bounds," submitted.
19. J. C. Doyle, "Guaranteed Margins for LQG Regulators," *IEEE Trans. Autom. Contr.*, Vol. AC-23, pp. 756-757, 1978.
20. D. S. Bernstein and S. W. Greeley, "Robust Controller Synthesis Using the Maximum Entropy Design Equations," *IEEE Trans. Autom. Contr.*, Vol. AC-31, pp. 362-364, 1986.
21. D. S. Bernstein and W. M. Haddad, "The Optimal Projection Equations with Petersen-Hellot Bounds: Robust Stability and Performance via Fixed-Order Dynamic Compensation for Systems with Structured Real-Valued Parameter Uncertainty," *IEEE Trans. Autom. Contr.*, Vol. AC-33, pp. 578-582, 1988.

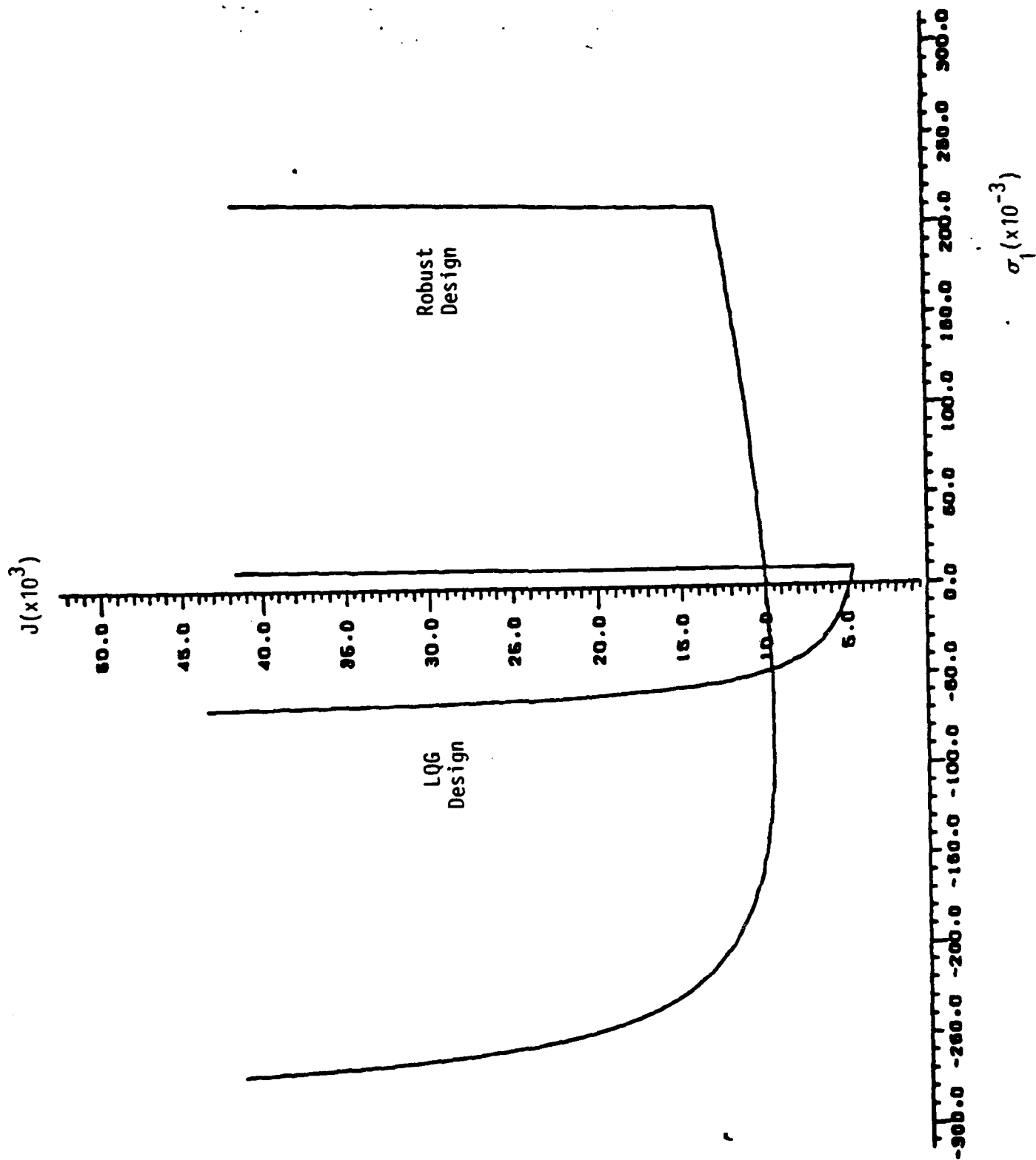


Figure 3

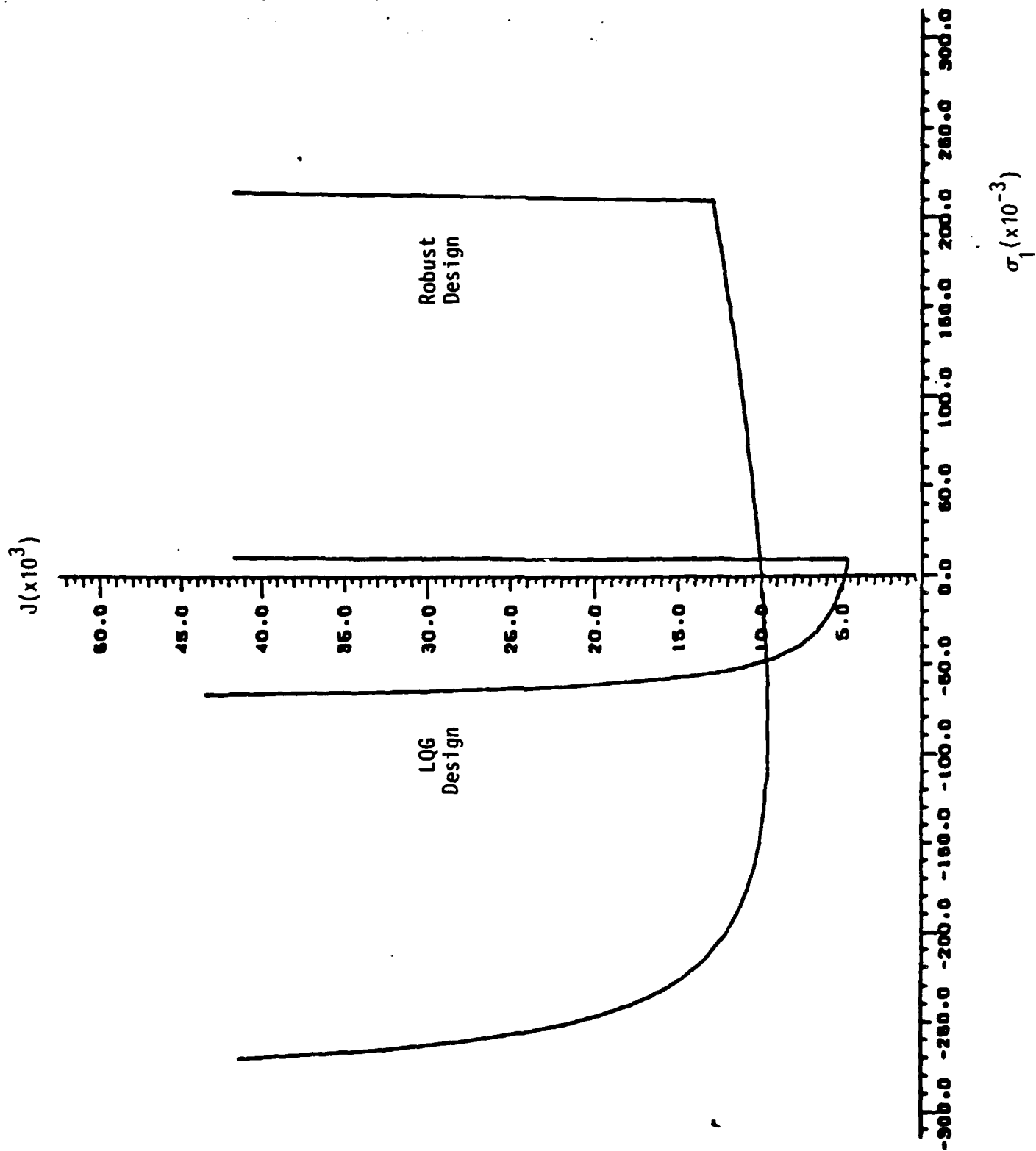


Figure 3

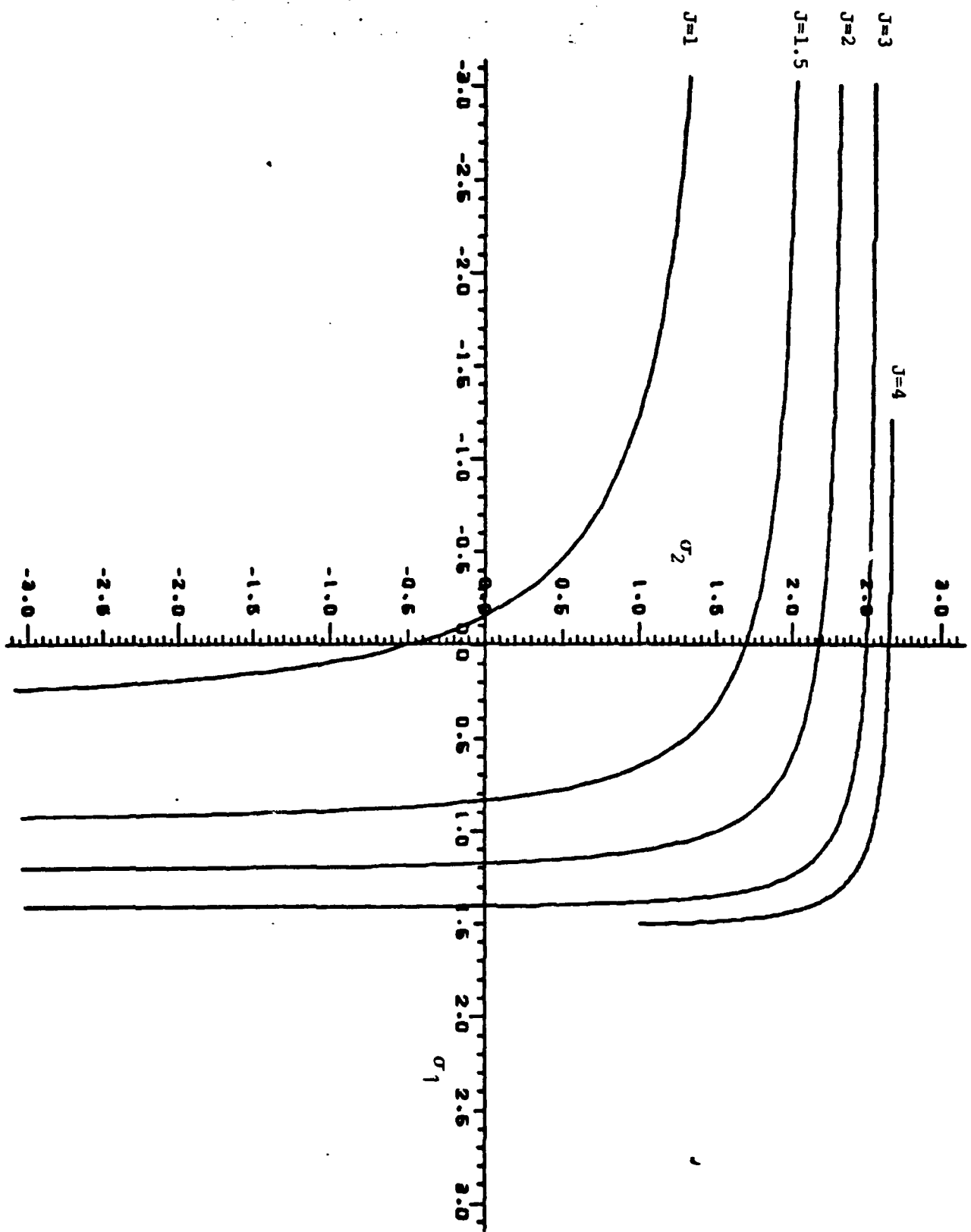


Figure 2

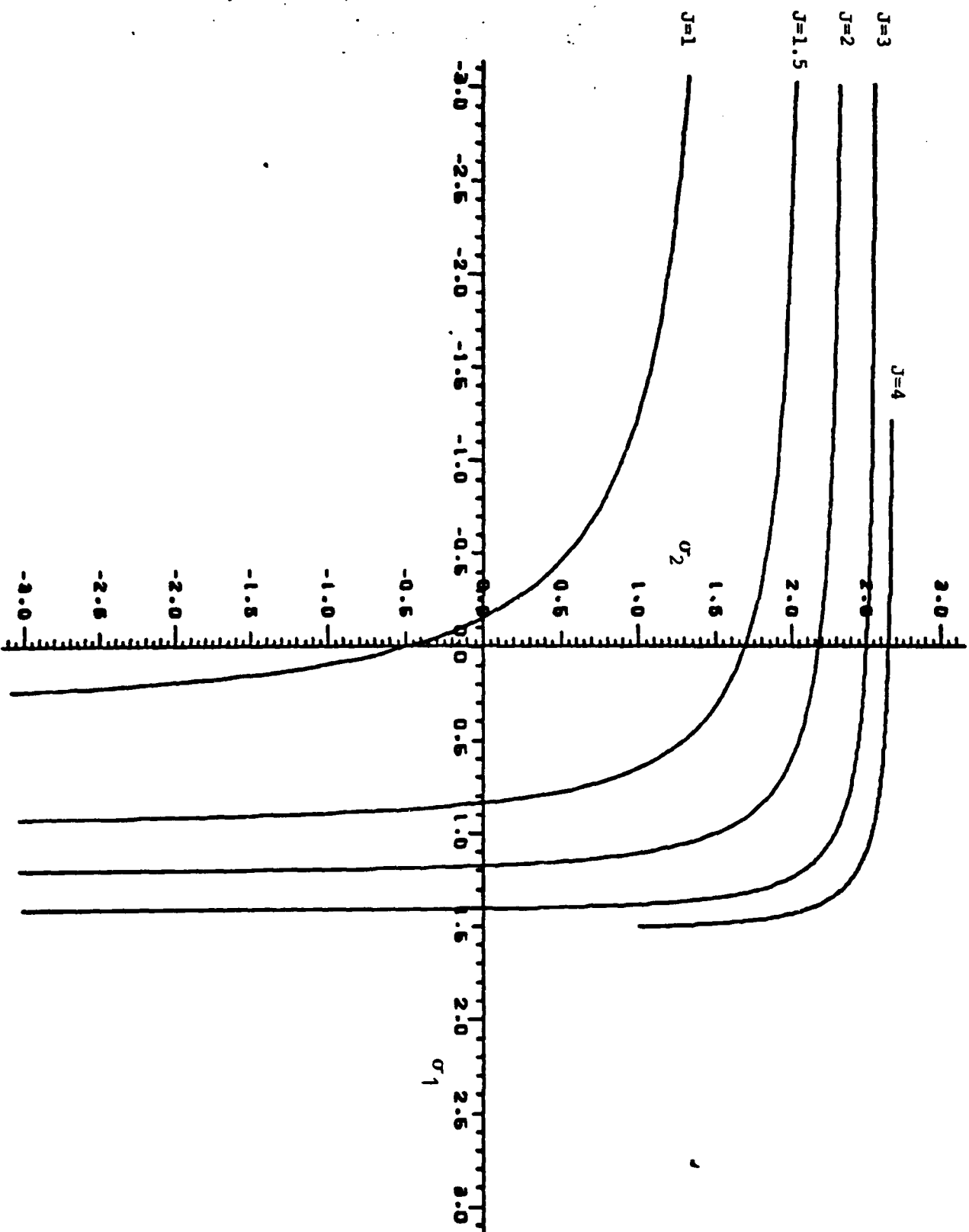


Figure 2

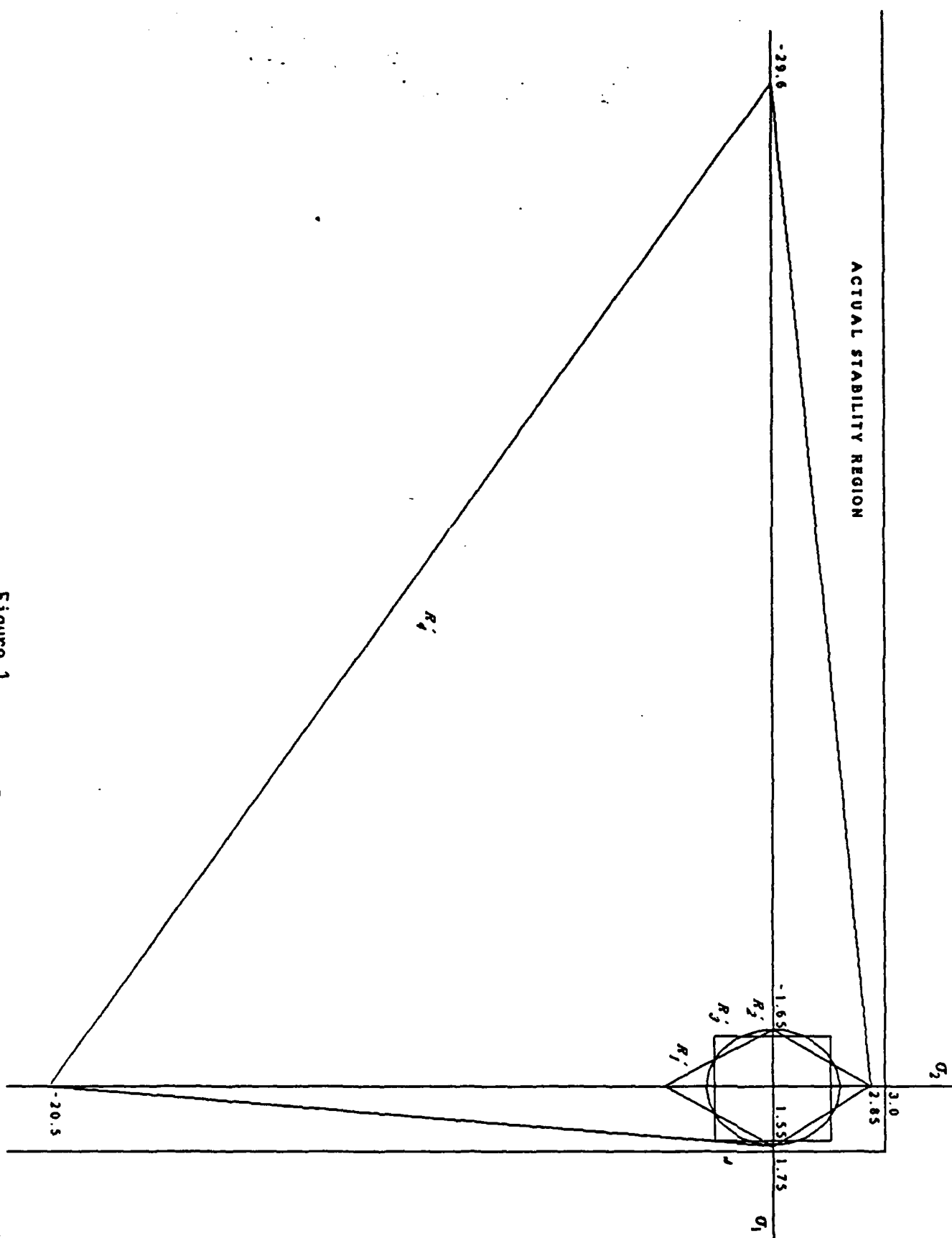


Figure 1

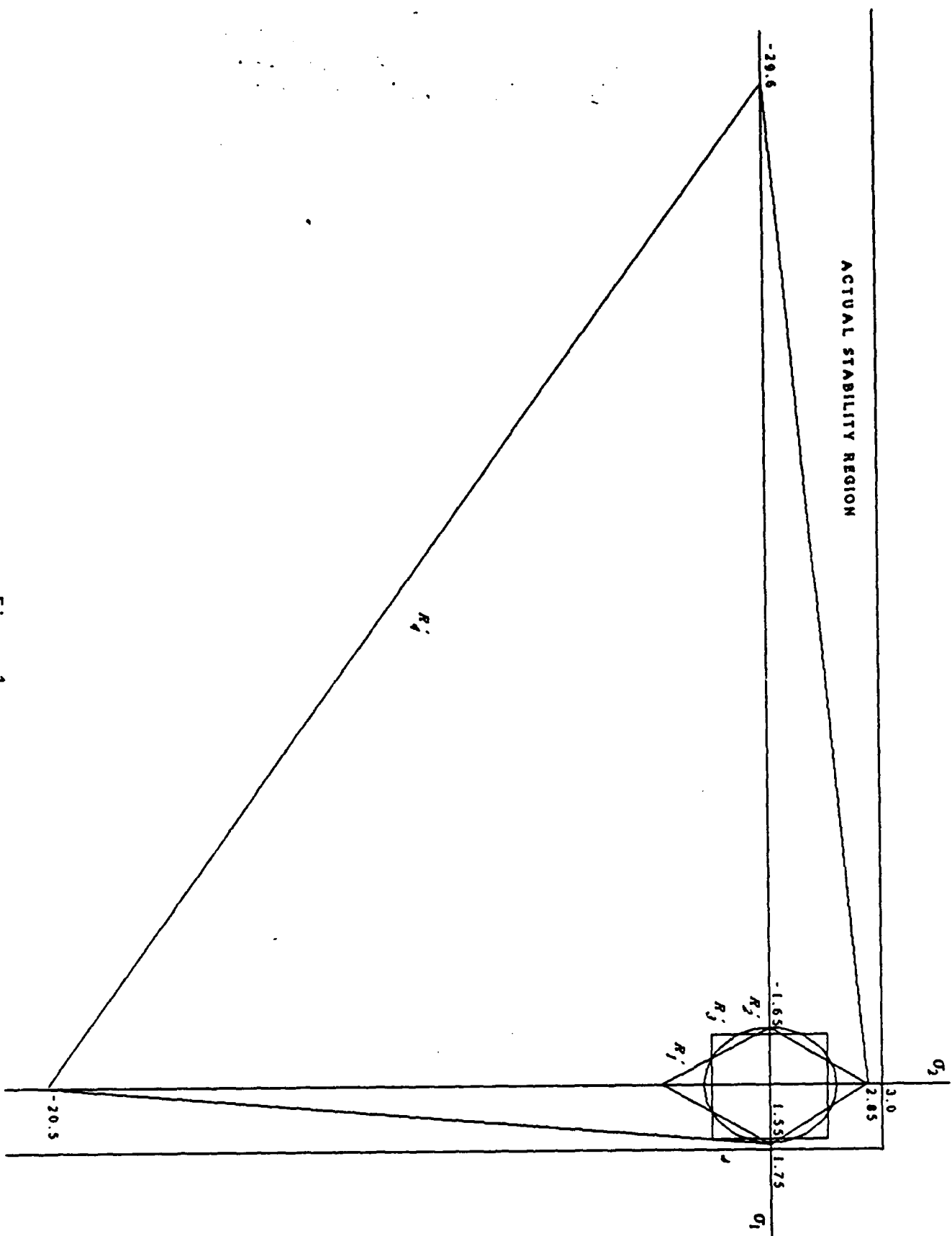


Figure 1

June 1988

**Robust Stability and Performance Analysis
for
State Space Systems
via
Quadratic Lyapunov Bounds**

by

Dennis S. Bernstein
Harris Corporation
Government Aerospace Systems Division
MS 22/4848
Melbourne, FL 32902

Wassim M. Haddad
Department of Mechanical and
Aerospace Engineering
Florida Institute of Technology
Melbourne, FL 32901

Abstract

For a given asymptotically stable linear dynamic system it is often of interest to determine whether stability is preserved as the system varies within a specified class of uncertainties. If, in addition, there also exist associated performance measures (such as the steady-state variances of selected state variables), it is desirable to assess the worst-case performance over a class of plant variations. These are problems of robust stability and performance analysis. In the present paper we consider quadratic Lyapunov bounds to obtain a simultaneous treatment of both robust stability and performance. The approach is based upon the construction of modified Lyapunov equations which provide sufficient conditions for robust stability along with robust performance bounds. One of the principal features of the paper is the unified treatment and extension of several quadratic Lyapunov bounds developed previously for feedback control design.

Key Words: robust analysis, stability, performance, Lyapunov equations, structured uncertainty

Supported in part by the Air Force Office of Scientific Research under contract F49620-86-C-0002.

1. Introduction

Unavoidable discrepancies between mathematical models and real-world systems can result in degradation of control-system performance including instability ([1,2]). Ideally, feedback control systems should be designed to be *robust* with respect to uncertainties, or perturbations, in the plant characteristics. Such uncertainties may arise either due to limitations in performing system identification prior to control-system implementation or because of unpredictable plant changes which occur during operation. Thus robustness *analysis* must play a key role in control-system design. That is, given an existing or proposed control system, determine the performance degradation due to variations in the plant.

In performing robustness analysis there are two principal concerns, namely, stability robustness and performance robustness. Stability robustness addresses the qualitative question as to whether or not the system remains stable for all plant perturbations within a specified class of uncertainties. A related problem involves determining the largest class of plant perturbations under which stability is preserved. Once robust stability has been ascertained, it is of interest to determine *quantitatively* the degradation of performance within a given robust stability range. In practice it is often desirable to determine the *worst case* performance as a measure of degradation.

The concern for both robust stability and performance can be traced back to the earliest developments in control theory. Design specifications such as gain and phase margin have traditionally been used to gauge system reliability in the face of uncertainty. In the modern control literature considerable effort has focused on rigorous robustness analysis and design techniques in a variety of settings. Analysis and synthesis results have been developed for both state-space and frequency-domain plant models to address structured parameter variations as well as normed-neighborhood uncertainty ([3-7]).

The present paper is concerned solely with the analysis of structured real-valued parameter uncertainty within the context of state space models. Motivation for such problems is clearly illustrated by simple examples given in [1,2]. These examples show that standard linear-quadratic methods used to design either full-state feedback controllers or dynamic compensators may result in closed-loop systems which are arbitrarily sensitive to structured real-valued plant parameter variations. A particularly effective technique for analyzing robust stability is to construct a Lyapunov function which guarantees stability of the system as the uncertain parameters vary over a specified range. Using the quadratic Lyapunov function $V(x) = x^T P x$ this technique has been extensively

developed for both analysis and synthesis (see, e.g., [8-37]).

Although both robust stability and performance are of interest in practice, most of the literature involving quadratic Lyapunov functions is confined to the problem of robust stability. A notable exception is the early work of Chang and Peng ([9]) which also provides bounds on worst-case quadratic performance within full-state feedback control design. In the present paper we further extend the approach of [9] to obtain a series of results for analyzing both robust stability and performance. As will be seen, these results also provide substantial unification of more recent results pertaining only to robust stability.

To illustrate the basis for our approach, consider the system

$$\dot{x}(t) = (A + \Delta A)x(t) + D_0 w(t), \quad (1.1)$$

where $x(t)$ is an n -vector, A is an $n \times n$ matrix denoting the nominal dynamics matrix, ΔA denotes an uncertain perturbation of A belonging to a specified set \mathcal{U} , and $D_0 w(t)$ is (for now) a white noise signal of specified intensity $V \triangleq D_0 D_0^T$. System (1.1) may, for example, denote a control system in closed-loop configuration. For the system (1.1) the performance involves the steady-state covariance of specified outputs $E_0 x(t)$. In practice the diagonal elements of the output covariance are measures of the ability of the external disturbances $D_0 w(t)$ to affect specified states. In the presence of uncertainties ΔA , it is of interest to determine the *worst case* steady-state values of selected state variances. Thus, we define the scalar performance criterion

$$J_S(\mathcal{U}) \triangleq \sup_{\Delta A \in \mathcal{U}} \limsup_{t \rightarrow \infty} \mathbb{E} \left\{ [E_0 x(t)]^T [E_0 x(t)] \right\}, \quad (1.2)$$

where "IE" denotes expectation. To evaluate (1.2) define the state covariance

$$Q(t) \triangleq \mathbb{E}[x(t)x^T(t)], \quad (1.3)$$

which satisfies the Lyapunov differential equation

$$\dot{Q}_{\Delta A}(t) = (A + \Delta A)Q_{\Delta A}(t) + Q_{\Delta A}(t)(A + \Delta A)^T + V, \quad (1.4)$$

so that (1.2) becomes

$$J_S(\mathcal{U}) = \sup_{\Delta A \in \mathcal{U}} \limsup_{t \rightarrow \infty} \text{tr } Q_{\Delta A}(t)R, \quad (1.5)$$

where $R \triangleq E_0 E_0^T$. To guarantee both robust stability and performance we consider modified algebraic Lyapunov equations of the form

$$0 = AQ + QA^T + \mathcal{N}(Q) + V \quad (1.6)$$

where $\Omega(\cdot)$ is a matrix operator satisfying

$$\Delta A Q + Q \Delta A^T \leq \Omega(Q) \quad (1.7)$$

for all $\Delta A \in \mathcal{U}$ and all nonnegative-definite matrices Q . The ordering in (1.7) is defined with respect to the cone of nonnegative-definite matrices. Our results are based on the following robust stability and performance result. If there exists a nonnegative-definite solution Q to (1.6) where $\Omega(\cdot)$ satisfies (1.7), then $A + \Delta A$ is asymptotically stable for all $\Delta A \in \mathcal{U}$ and, furthermore,

$$J_S(\mathcal{U}) \leq \text{tr } QR. \quad (1.8)$$

The performance bound (1.8) follows from the fact that since $A + \Delta A$ is asymptotically stable, $Q_{\Delta A} \triangleq \lim_{t \rightarrow \infty} Q_{\Delta A}(t)$ exists and satisfies

$$0 = (A + \Delta A)Q_{\Delta A} + Q_{\Delta A}(A + \Delta A)^T + V. \quad (1.9)$$

Now subtracting (1.9) from (1.6) implies

$$Q_{\Delta A} \leq Q, \quad (1.10)$$

which with (1.5) yields (1.8).

Since the ordering induced by the cone of nonnegative-definite matrices is only a partial ordering, it should not be expected that there exists an operator $\Omega(\cdot)$ satisfying (1.7) which is a least upper bound. Indeed, connections between the result outlined above and the approach of [9] as well as more recent work arise from alternative definitions of the operator $\Omega(\cdot)$. To illustrate these connections assume for convenience that ΔA is of the form

$$\Delta A = \sigma_1 A_1, \quad |\sigma_1| \leq \delta_1, \quad (1.11)$$

where σ_1 is an uncertain real scalar parameter assumed only to satisfy the stated bounds, and A_1 is a known matrix denoting the structure of the parametric uncertainty. The original definition of $\Omega(\cdot)$ in [9] was given by

$$\Omega(Q) = \delta_1 |A_1 Q + Q A_1^T|, \quad (1.12)$$

where $|\cdot|$ denotes the nonnegative-definite matrix obtained by replacing each eigenvalue by its absolute value. This bound was studied in [9,12] for full-state feedback design. More recently, the quadratic bound

$$\Omega(Q) = \delta_1 [D + Q E Q] \quad (1.13)$$

has been considered, where $D \triangleq D_1 D_1^T$, $E \triangleq E_1^T E_1$, and D_1, E_1 are a factorization of A_1 of the form $A_1 = D_1 E_1$. Bound (1.13) was studied in [29] for robustness analysis and in [17,25,28,30,33,36] for robust controller synthesis. A third bound which has also been considered is the linear bound

$$\Omega(Q) = \delta_1[\alpha Q + \alpha^{-1} A_1 Q A_1^T], \quad (1.14)$$

where α is an arbitrary positive scalar. As shown in [33], bound (1.14) arises from a multiplicative white noise model with exponential disturbance weighting. Control-design applications of bound (1.14) are given in [23,27,33,34,35].

The principal contribution of the present paper is thus a unified development of bounds (1.12)–(1.14) for both robust stability and performance analysis. In addition, we present a systematic approach which pays careful attention to the structure of the uncertainty set \mathcal{U} . For example, we show that bound (1.12) guarantees stability over a rectangular uncertainty set while (1.14) is most naturally associated with an ellipsoidal region. Furthermore, to provide a methodical development, we consider three classes of bounds (Type I, II and III) which operate by exploiting, respectively, the symmetry of $\Delta A Q + Q \Delta A^T$, the structure of Q , and the structure of ΔA . This approach clarifies the relationships among different bounds and suggests several new bounds.

Finally, the present paper also considers an alternative functional for robust performance analysis. Specifically, in place of white noise disturbances, we reinterpret $w(t)$ in (1.1) as a deterministic L_2 signal as in H_∞ theory ([6]). By imposing an L_∞ norm on the output $E_0 x(t)$, the corresponding performance measure is given by (see [38])

$$J_D(\mathcal{U}) = \sup_{\Delta A \in \mathcal{U}} \limsup_{t \rightarrow \infty} \lambda_{\max}(Q \Delta A(t) R), \quad (1.15)$$

in contrast to (1.5). Both performance measures $J_S(\mathcal{U})$ and $J_D(\mathcal{U})$ are considered in the paper.

The contents of the paper are as follows. After summarizing notation later in this section, the Robust Stability Problem, Stochastic Robust Performance Problem, and Deterministic Robust Performance Problem are introduced in Section 2. In Section 3 the basic result guaranteeing robust stability and performance (Theorem 3.1) is stated. This result is easily stated and forms the basis for all later developments. A dual version of Theorem 3.1 (Theorem 4.1) provides additional sufficient conditions and clarifies connections to traditional robust stability results. The bound $\Omega(\cdot)$ and its dual $\Lambda(\cdot)$ are given concrete forms in Section 5. In Section 6, the bounds of Section 5 are merged with Theorem 3.1 to yield the main results guaranteeing robust stability and performance (Theorems

6.1-6.5) via modified Lyapunov equations. In Section 7 we analyze the modified Lyapunov equations with regard to existence, uniqueness, and monotonicity of solutions. Additional bounds are derived in Section 8 by utilizing a recursive substitution technique, while both upper and lower bounds are obtained in Section 9. Finally, illustrative numerical examples are considered in Sections 10 and 11.

Notation

Note: All matrices have real entries

| | |
|---|---|
| $\mathbb{R}, \mathbb{R}^{r \times s}, \mathbb{R}^r, \mathbb{E}$ | real numbers, $r \times s$ real matrices, $\mathbb{R}^{r \times 1}$, expectation |
| I_r | $r \times r$ identity matrix |
| asymptotically stable matrix | matrix with eigenvalues in open left half plane |
| \mathbb{S}^r | $r \times r$ symmetric matrices |
| \mathbb{N}^r | $r \times r$ symmetric nonnegative-definite matrices |
| \mathbb{P}^r | $r \times r$ symmetric positive-definite matrices |
| $Z_1 \geq Z_2$ | $Z_1 - Z_2 \in \mathbb{N}^r, \quad Z_1, Z_2 \in \mathbb{S}^r$ |
| $Z_1 > Z_2$ | $Z_1 - Z_2 \in \mathbb{P}^r, \quad Z_1, Z_2 \in \mathbb{S}^r$ |
| $\text{tr } Z, Z^T$ | trace of Z , transpose of Z |
| $\lambda_i(Z)$ | eigenvalue of matrix Z |
| $\lambda_{\max}(Z)$ | maximum eigenvalue of matrix Z having real spectrum |

2. Robust Stability and Performance Problems

Let $\mathcal{U} \subset \mathbb{R}^{n \times n}$ denote a set of perturbations ΔA of a given nominal dynamics matrix $A \in \mathbb{R}^{n \times n}$. Throughout the paper it is assumed that A is asymptotically stable and that $0 \in \mathcal{U}$. We begin by considering the question of whether or not $A + \Delta A$ is asymptotically stable for all $\Delta A \in \mathcal{U}$.

Robust Stability Problem. Determine whether the linear system

$$\dot{x}(t) = (A + \Delta A)x(t), \quad t \in [0, \infty), \quad (2.1)$$

is asymptotically stable for all $\Delta A \in \mathcal{U}$.

To consider the problem of robust performance it is necessary to introduce external disturbances. In this paper we consider both stochastic and deterministic disturbance models. The

stochastic disturbance model involves white noise signals as in standard LQG theory while the deterministic disturbance model involves L_2 signals as in H_∞ theory ([6]). By defining an appropriate performance measure for each disturbance class it turns out that we can provide a simultaneous treatment of both cases.

We first consider the case of stochastic disturbances. In this case the robust performance problem concerns the worst-case magnitude of the expected value of a quadratic form $[E_0 x(t)]^T [E_0 x(t)]$, where the matrix $E_0 \in \mathbb{R}^{q \times n}$ defines the output states, when the system is subjected to a standard white noise disturbance $w(t) \in \mathbb{R}^d$ with weighting $D_0 \in \mathbb{R}^{n \times d}$.

Stochastic Robust Performance Problem. For the disturbed linear system

$$\dot{x}(t) = (A + \Delta A)x(t) + D_0 w(t), \quad t \in [0, \infty), \quad (2.2)$$

where $w(\cdot)$ is a d -dimensional white noise signal with intensity I_d , determine a performance bound β_s satisfying

$$J_{RS}(\mathcal{U}) \triangleq \sup_{\Delta A \in \mathcal{U}} \limsup_{t \rightarrow \infty} \mathbb{E} \left\{ [E_0 x(t)]^T [E_0 x(t)] \right\} \leq \beta_s. \quad (2.3)$$

The system (2.2) may denote, for example, a control system in closed-loop configuration subjected to external white noise disturbances for which $z(t) \triangleq E_0 x(t)$ may be the state regulation error. Such specializations are not required for this development, however.

Of course, since E_0 and D_0 may be rank deficient, there may be cases in which a finite performance bound β_s satisfying (2.3) exists while (2.1) is not asymptotically stable over \mathcal{U} . In practice, however, robust performance is mainly of interest when (2.1) is robustly stable. In this case the performance $J_S(\mathcal{U})$ is given in terms of the steady-state second moment of the state. The following result from linear system theory will be useful. For convenience define the $n \times n$ nonnegative-definite matrices

$$R \triangleq E_0^T E_0, \quad V \triangleq D_0 D_0^T. \quad (2.4)$$

Lemma 2.1. Suppose (2.1) is asymptotically stable for all $\Delta A \in \mathcal{U}$. Then

$$J_S(\mathcal{U}) = \sup_{\Delta A \in \mathcal{U}} \text{tr } Q_{\Delta A} R, \quad (2.5)$$

where $n \times n$ $Q_{\Delta A} \triangleq \lim_{t \rightarrow \infty} \mathbb{E}[x(t)x^T(t)]$ is given by

$$Q_{\Delta A} = \int_0^\infty e^{(A+\Delta A)t} V e^{(A+\Delta A)^T t} dt, \quad (2.6)$$

which is the unique, nonnegative-definite solution to

$$0 = (A + \Delta A)Q_{\Delta A} + Q_{\Delta A}(A + \Delta A)^T + V. \quad (2.7)$$

To state the Deterministic Robust Performance Problem some additional notation is required. For a measurable function $w : [0, \infty) \rightarrow \mathbb{R}^d$ define

$$\|w(\cdot)\|_2 \triangleq \left\{ \int_0^\infty w^T(t)w(t)dt \right\}^{\frac{1}{2}}. \quad (2.8)$$

Note that definition (2.8) is an L_2 function norm with a Euclidean spatial norm. We now reconsider (2.2) with $w(\cdot)$ now interpreted as a square-integrable function. In this case the robust performance problem concerns the worst-case L_∞ norm of a quadratic form $[E_0 x(t)]^T [E_0 x(t)]$.

Deterministic Robust Performance Problem. For the disturbed linear system (2.2), where $\|w(\cdot)\|_2 \leq 1$, determine a performance bound β_D satisfying

$$J_D(\mathcal{U}) \triangleq \sup_{\Delta A \in \mathcal{U}} \sup_{\|w(\cdot)\|_2 \leq 1} \sup_{t \in [0, \infty)} \left\{ [E_0 x(t)]^T [E_0 x(t)] \right\} \leq \beta_D. \quad (2.9)$$

The following result is proved in [38].

Lemma 2.2. Suppose (2.1) is asymptotically stable for all $\Delta A \in \mathcal{U}$. Then

$$J_D(\mathcal{U}) = \sup_{\Delta A \in \mathcal{U}} \lambda_{\max}(Q_{\Delta A} R), \quad (2.10)$$

where $Q_{\Delta A}$ is the unique, nonnegative-definite solution to (2.7).

Remark 2.1. Although $J_S(\mathcal{U})$ and $J_D(\mathcal{U})$ arise from different mathematical settings they are quite similar in form. Note that in general $J_D(\mathcal{U}) \leq J_S(\mathcal{U})$, and $J_D(\mathcal{U}) = J_S(\mathcal{U})$ if $\text{rank } R = 1$.

Remark 2.2. In Lemma 2.2 $Q_{\Delta A}$ can be viewed as the controllability Gramian for the pair $(A + \Delta A, D_0)$ rather than the state covariance.

In the present paper our approach is to obtain robust stability as a consequence of sufficient conditions for robust performance. Such conditions are developed in the following sections.

3. Sufficient Conditions for Robust Stability and Performance

The key step in obtaining robust stability and performance is to bound the uncertain terms $\Delta A Q + Q \Delta A^T$ in the Lyapunov equation (2.7) by means of a function $\Omega(Q)$. The nonnegative-definite solution Q of this modified Lyapunov equation is then guaranteed to be an upper bound

for $Q_{\Delta A}$. The following easily proved result is fundamental and forms the basis for all later developments. The hypothesis of this result are of a general nature and are not intended to be directly verifiable. Suitably verifiable hypotheses are discussed later.

Theorem 3.1. Let $\Omega : \mathbb{R}^n \rightarrow \mathbb{R}^n$ be such that

$$\Delta A Q + Q \Delta A^T \leq \Omega(Q), \quad \Delta A \in \mathcal{U}, \quad Q \in \mathbb{R}^n, \quad (3.1)$$

and suppose there exists $Q \in \mathbb{R}^n$ satisfying

$$0 = A Q + Q A^T + \Omega(Q) + V. \quad (3.2)$$

Then

$$(A + \Delta A, [V + \Omega(Q) - (\Delta A Q + Q \Delta A^T)]^{\frac{1}{2}}) \text{ is stabilizable, } \Delta A \in \mathcal{U}, \quad (3.3)$$

if and only if

$$A + \Delta A \text{ is asymptotically stable, } \Delta A \in \mathcal{U}. \quad (3.4)$$

In this case,

$$Q_{\Delta A} \leq Q, \quad \Delta A \in \mathcal{U}, \quad (3.5)$$

where $Q_{\Delta A} \in \mathbb{R}^n$ is given by (2.7), and

$$J_S(\mathcal{U}) \leq \text{tr } Q R, \quad (3.6)$$

$$J_D(\mathcal{U}) \leq \lambda_{\max}(Q R). \quad (3.7)$$

In addition, there exists $\Delta A \in \mathcal{U}$ such that $(A + \Delta A, [V + \Omega(Q) - (\Delta A Q + Q \Delta A^T)]^{\frac{1}{2}})$ is controllable if and only if Q is positive definite. In this case $(A + \Delta A, [V + \Omega(Q) - (\Delta A Q + Q \Delta A^T)]^{\frac{1}{2}})$ is controllable for all $\Delta A \in \mathcal{U}$.

Proof. We stress that in (3.1) " Q " denotes an arbitrary element of \mathbb{R}^n while Q in (3.2) denotes a specific solution of the modified Lyapunov equation. This minor abuse of notation considerably simplifies the presentation. Now note that for all $\Delta A \in \mathbb{R}^{n \times n}$ (3.2) is equivalent to

$$0 = (A + \Delta A) Q + Q (A + \Delta A)^T + \Omega(Q) - (\Delta A Q + Q \Delta A^T) + V. \quad (3.8)$$

Hence, by assumption (3.8) has a solution $Q \in \mathbb{R}^n$ for all $\Delta A \in \mathbb{R}^{n \times n}$. If ΔA is restricted to the set \mathcal{U} then by (3.1) $\Omega(Q) - (\Delta A Q + Q \Delta A^T)$ is nonnegative definite. Now if the stabilizability condition (3.3) holds for all $\Delta A \in \mathcal{U}$, it follows from Lemma 12.2 of [39] that $A + \Delta A$ is asymptotically stable

for all $\Delta A \in \mathcal{U}$. Conversely, if $A + \Delta A$ is asymptotically stable for all $\Delta A \in \mathcal{U}$, then (3.3) holds. Next, subtracting (2.7) from (3.8) yields

$$0 = (A + \Delta A)(Q - Q_{\Delta A}) + (Q - Q_{\Delta A})(A + \Delta A)^T + \Omega(Q) - (\Delta A Q + Q \Delta A^T), \quad \Delta A \in \mathcal{U},$$

or, equivalently, (since $A + \Delta A$ is asymptotically stable for all $\Delta A \in \mathcal{U}$)

$$Q - Q_{\Delta A} = \int_0^\infty e^{(A + \Delta A)t} [\Omega(Q) - (\Delta A Q + Q \Delta A^T)] e^{(A + \Delta A)^T t} dt \geq 0, \quad \Delta A \in \mathcal{U},$$

which implies (3.5). The performance bound (3.6) is now an immediate consequence of (3.5). To prove (3.7) we note that it follows from Corollary 7.7.4 of [40] that if $0 \leq M_1 \leq M_2$ then $\lambda_{\max}(M_1) \leq \lambda_{\max}(M_2)$. Thus

$$J_D(\mathcal{U}) = \lambda_{\max}(Q_{\Delta A} R) = \lambda_{\max}(E_0 Q_{\Delta A} E_0^T) \leq \lambda_{\max}(E_0 Q E_0^T) = \lambda_{\max}(Q R). \quad (3.9)$$

Finally, it follows from (3.8) that the controllability condition holds for some $\Delta A \in \mathcal{U}$ if and only if the Gramian Q is positive definite. Since Q is also the Gramian corresponding to $A + \Delta A$ for all $\Delta A \in \mathcal{U}$, then controllability holds for all $\Delta A \in \mathcal{U}$. \square

For convenience we shall say that $\Omega(\cdot)$ *bounds* \mathcal{U} if (3.1) is satisfied. To apply Theorem 3.1, one first specifies a function $\Omega(\cdot)$ and an uncertainty set \mathcal{U} such that $\Omega(\cdot)$ bounds \mathcal{U} . If the existence of a nonnegative-definite solution Q to (3.2) satisfying (3.3) or, equivalently, (3.4) can be determined analytically or numerically, then robust stability is guaranteed. One can then enlarge \mathcal{U} , modify $\Omega(\cdot)$, and again attempt to solve (3.2). If, however, a nonnegative-definite solution to (3.2) cannot be determined, then \mathcal{U} must be decreased in size until (3.2) is solvable. For example, $\Omega(\cdot)$ can be replaced by $\varepsilon \Omega(\cdot)$ to bound $\varepsilon \mathcal{U}$, where $\varepsilon > 1$ enlarges \mathcal{U} and $\varepsilon < 1$ shrinks \mathcal{U} . Of course, the actual range of uncertainty which can be bounded depends upon the nominal matrix A , the function $\Omega(\cdot)$, and the structure of \mathcal{U} . In Section 5 the uncertainty set \mathcal{U} and bound $\Omega(\cdot)$ satisfying (3.1) are given concrete forms. We complete this section with several observations.

Remark 3.1. In applying Theorem 3.1 it may be convenient to replace condition (3.3) with a stronger condition which is easier to verify in practice. Clearly, (3.3) is satisfied if $V + \Omega(Q) - (\Delta A Q + Q \Delta A^T)$ is positive definite for all $\Delta A \in \mathcal{U}$. This will be the case, for example, if either V is positive definite or strict inequality holds in (3.1). Also, it follows from Theorem 3.6 of [39] that (3.3) is implied by the stronger condition

$$(A + \Delta A, V^{\frac{1}{2}}) \text{ is stabilizable}, \quad \Delta A \in \mathcal{U}. \quad (3.10)$$

Similar remarks apply to the controllability condition.

Remark 3.2. If only robust stability is of interest, then the noise intensity V need not have physical significance. In this case one may either set $V = \epsilon I_n$ where $\epsilon > 0$ is small to satisfy $V > 0$, or set $V = 0$ and require that strict inequality hold in (3.1).

Remark 3.3. Since A is stable Q satisfying (3.2) is given by

$$Q = \int_0^\infty e^{At} [\Omega(Q) + V] e^{A^T t} dt, \quad (3.11)$$

or, equivalently,

$$Q = \int_0^\infty e^{At} \Omega(Q) e^{A^T t} dt + Q_0, \quad (3.12)$$

where $Q_0 \in \mathbb{R}^n$ is defined by

$$Q_0 \triangleq \int_0^\infty e^{At} V e^{A^T t} dt \quad (3.13)$$

and satisfies

$$0 = A Q_0 + Q_0 A^T + V. \quad (3.14)$$

Note that $Q_0 \leq Q$ and that the nominal performances $J_S(\{0\})$ and $J_D(\{0\})$ are given by $\text{tr } Q_0 R$ and $\lambda_{\max}(Q_0 R)$, respectively.

Remark 3.4. Using (3.11) it is also useful to note that the bound for $J_S(\mathcal{U})$ given by (3.6) can be written as

$$\text{tr } QR = \text{tr} \int_0^\infty e^{At} [\Omega(Q) + V] e^{A^T t} dt R = \text{tr } P_0 [\Omega(Q) + V], \quad (3.15)$$

where $P_0 \in \mathbb{R}^n$ is defined by

$$P_0 \triangleq \int_0^\infty e^{A^T t} R e^{At} dt \quad (3.16)$$

and satisfies

$$0 = A^T P_0 + P_0 A + R. \quad (3.17)$$

The bound $\text{tr } P_0 [\Omega(Q) + V]$ can be viewed as a dual formulation of the bound $\text{tr } QR$ since the roles of A and A^T are reversed. Dual bounds are developed in the following section. Note that $\text{tr } Q_0 R = \text{tr } P_0 V$.

Remark 3.5. If $\Omega(\cdot)$ bounds \mathcal{U} then clearly $\Omega(\cdot)$ bounds the convex hull of \mathcal{U} . Hence, only convex uncertainty sets \mathcal{U} need be considered. Next, we shall later use the obvious fact that if $\Omega'(\cdot)$ bounds \mathcal{U}' and $\Omega''(\cdot)$ bounds \mathcal{U}'' , then $\Omega'(\cdot) + \Omega''(\cdot)$ bounds $\mathcal{U}' + \mathcal{U}''$. Hence if \mathcal{U} can be decomposed

additively then it suffices to bound each component separately. Finally, if $\Omega(\cdot)$ bounds \mathcal{U} and there exists $\Omega' : \mathbb{R}^n \rightarrow \mathbb{R}^n$ such that $\Omega(Q) \leq \Omega'(Q)$ for all $Q \in \mathbb{R}^n$, then $\Omega'(\cdot)$ also bounds \mathcal{U} . That is, any *overbound* $\Omega'(\cdot)$ for $\Omega(\cdot)$ also bounds \mathcal{U} . Of course, as we shall see, it is quite possible that an overbound $\Omega'(\cdot)$ for $\Omega(\cdot)$ may bound a set \mathcal{U}' which is larger than the "original" uncertainty set \mathcal{U} .

4. Dual Sufficient Conditions for Robust Stability and Performance

As noted in Remark 3.4, the performance bound $\text{tr } QR$ given by (3.6) can be expressed equivalently in terms of a dual variable P_0 for which the roles of A and A^T are reversed. Using a similar technique, additional conditions for robust stability and performance can be obtained by developing a dual version of Theorem 3.1. A prime motivation for developing such dual bounds is to draw connections with previous results in the literature relating to robust stability. In particular, note that traditional robust stability techniques based upon the Lyapunov function $V(x) = x^T Px$ lead to dual conditions. Robust performance bounds within the dual formulation are difficult to motivate without first developing the primal performance bounds. In addition, the dual bounds may, for certain problems, yield larger stability regions and sharper performance bounds than the "primal" bounds introduced in the previous section.

Lemma 4.1. Suppose (2.1) is asymptotically stable for all $\Delta A \in \mathcal{U}$. Then

$$J_S(\mathcal{U}) = \sup_{\Delta A \in \mathcal{U}} \text{tr } P_{\Delta A} V, \quad (4.1)$$

where $n \times n$ $P_{\Delta A}$ is the unique, nonnegative-definite solution to

$$0 = (A + \Delta A)^T P_{\Delta A} + P_{\Delta A} (A + \Delta A) + R. \quad (4.2)$$

Proof. It need only be noted that

$$\text{tr } Q_{\Delta A} R = \text{tr} \int_0^\infty e^{(A+\Delta A)t} V e^{(A+\Delta A)^T t} dt R = \text{tr } P_{\Delta A} V,$$

where

$$P_{\Delta A} \triangleq \int_0^\infty e^{(A+\Delta A)^T t} R e^{(A+\Delta A)t} dt$$

satisfies (4.2). \square

The proof of Lemma 4.1 relied on the fact that $\text{tr } Q_{\Delta A} R = \text{tr } P_{\Delta A} V$. However, it is not necessarily true that $\lambda_{\max}(Q_{\Delta A} R) = \lambda_{\max}(P_{\Delta A} V)$ even when $\Delta A = 0$. For example, if $A =$

$\begin{bmatrix} -1 & 0 \\ 0 & -2 \end{bmatrix}$, $R = I_2$, and $V = \begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix}$ then $Q_0 R = \begin{bmatrix} 1 & 1/3 \\ 1/3 & 1/4 \end{bmatrix}$ and $P_0 V = \begin{bmatrix} 1 & 1/2 \\ 1/4 & 1/4 \end{bmatrix}$ and thus $\lambda_{\max}(Q_0 R) = (15 + \sqrt{145})/24$ and $\lambda_{\max}(P_0 V) = (5 + \sqrt{17})/8$. Thus to obtain a suitable dual version of $J_D(\mathcal{U})$ we need to define a dual deterministic cost $\hat{J}_D(\mathcal{U})$. This can be done if the disturbance signals are taken to be integrable rather than square integrable. Thus, for measurable $w : [0, \infty) \rightarrow \mathbb{R}^d$ define

$$\|w(\cdot)\|_1 \triangleq \int_0^\infty [w^T(t)w(t)]^{\frac{1}{2}} dt, \quad (4.3)$$

which is an L_1 function norm with a Euclidean spatial norm. The dual deterministic cost $\hat{J}_D(\mathcal{U})$ is thus defined by

$$\hat{J}_D(\mathcal{U}) \triangleq \sup_{\Delta A \in \mathcal{U}} \sup_{\|w(\cdot)\|_1 \leq 1} \|Ex(\cdot)\|_2^2, \quad (4.4)$$

where $Ex(\cdot)$ is measured according to the energy norm (2.8). The following dual result can also be found in [38].

Lemma 4.2. Suppose (2.1) is asymptotically stable for all $\Delta A \in \mathcal{U}$. Then

$$\hat{J}_D(\mathcal{U}) = \lambda_{\max}(P_{\Delta A} V), \quad (4.5)$$

where $P_{\Delta A}$ is the unique, nonnegative-definite solution to (4.2).

The dual version of Theorem 3.1 can now be stated.

Theorem 4.1. Let $\Lambda : \mathbb{N}^n \rightarrow \mathbb{N}^n$ be such that

$$\Delta A^T P + P \Delta A \leq \Lambda(P), \quad \Delta A \in \mathcal{U}, \quad P \in \mathbb{N}^n, \quad (4.6)$$

and suppose there exists $P \in \mathbb{N}^n$ satisfying

$$0 = A^T P + P A + \Lambda(P) + R. \quad (4.7)$$

Then

$$\left([R + \Lambda(P) - (\Delta A^T P + P \Delta A)]^{\frac{1}{2}}, A + \Delta A \right) \text{ is detectable, } \Delta A \in \mathcal{U}, \quad (4.8)$$

if and only if

$$A + \Delta A \text{ is asymptotically stable, } \Delta A \in \mathcal{U}. \quad (4.9)$$

In this case,

$$P_{\Delta A} \leq P, \quad \Delta A \in \mathcal{U}, \quad (4.10)$$

where $P_{\Delta A}$ is given by (4.2), and

$$J_S(\mathcal{U}) \leq \text{tr } P V, \quad (4.11)$$

$$\hat{J}_D(\mathcal{U}) \leq \lambda_{\max}(PV). \quad (4.12)$$

In addition, there exists $\Delta A \in \mathcal{U}$ such that $\left([R + \Lambda(P) - (\Delta A^T P + P \Delta A)]^{\frac{1}{2}}, A + \Delta A\right)$ is observable if and only if P is positive definite. In this case $\left([R + \Lambda(P) - (\Delta A^T P + P \Delta A)]^{\frac{1}{2}}, A + \Delta A\right)$ is observable for all $\Delta A \in \mathcal{U}$.

Proof. The proof is completely analogous to the proof of Theorem 3.1. \square

Remark 4.1. Note that $\hat{J}_D(\mathcal{U}) \leq J_S(\mathcal{U})$ and $\hat{J}_D(\mathcal{U}) = J_S(\mathcal{U})$ if $\text{rank } V = 1$. Combining this fact with Remark 2.1, it follows that $J_D(\mathcal{U}) = \hat{J}_D(\mathcal{U})$ if both $\text{rank } R = 1$ and $\text{rank } V = 1$.

It is quite possible that the bounds $\text{tr } QR$ and $\text{tr } PV$ for $J_S(\mathcal{U})$ given by (3.6) and (4.11) may be different in spite of the fact as shown in the proof of Lemma 4.1 that $\text{tr } Q_{\Delta A} R = \text{tr } P_{\Delta A} V$. That is, depending upon $\Omega(\cdot)$ and $\Lambda(\cdot)$ either bound (3.6) or bound (4.11) may be better for a particular problem. In general, we have the following result.

Proposition 4.1. Let $\Omega(\cdot)$, $\Lambda(\cdot)$, Q , and P be as in Theorems 3.1 and 4.1, and let Q_0 and P_0 be given by (3.13) and (3.16), respectively. Then

$$\text{tr } Q_0 \Lambda(P) < \text{tr } P_0 \Omega(Q) \iff \text{tr } QR > \text{tr } PV, \quad (4.13)$$

$$\text{tr } Q_0 \Lambda(P) = \text{tr } P_0 \Omega(Q) \iff \text{tr } QR = \text{tr } PV, \quad (4.14)$$

$$\text{tr } Q_0 \Lambda(P) > \text{tr } P_0 \Omega(Q) \iff \text{tr } QR < \text{tr } PV. \quad (4.15)$$

Proof. Note that

$$\text{tr } QR = \int_0^\infty e^{A^T t} [\Omega(Q) + V] e^{A^T t} dt R = \text{tr } P_0 \Omega(Q) + \text{tr } \int_0^\infty e^{A^T t} V e^{A^T t} dt R$$

and

$$\text{tr } PV = \text{tr } \int_0^\infty e^{A^T t} [\Lambda(P) + R] e^{A^T t} dt V = \text{tr } Q_0 \Lambda(P) + \text{tr } \int_0^\infty e^{A^T t} R e^{A^T t} dt V$$

so that

$$\text{tr } QR - \text{tr } PV = \text{tr } P_0 \Omega(Q) - \text{tr } Q_0 \Lambda(P),$$

which yields (4.13)–(4.15). \square

Remark 4.2. Finally, as shown by example above, it is not generally true that $J_D(\mathcal{U}) = \hat{J}_D(\mathcal{U})$. Similarly, we should not expect that the bounds $\lambda_{\max}(QR)$ and $\lambda_{\max}(PV)$ for $J_D(\mathcal{U})$ and $\hat{J}_D(\mathcal{U})$ given by (3.7) and (4.12) are equal.

5. Construction of the Bounds $\Omega(\cdot)$ and $\Lambda(\cdot)$

As discussed in Section 1, we consider three distinct classes of bounds $\Omega(\cdot)$ denoted by Type I, Type II, and Type III. Roughly speaking, these bounds exploit, respectively, the symmetry of the Lyapunov terms $\Delta A Q + Q \Delta A^T$, the structure of Q , and the structure of ΔA . The dual bounds $\Lambda(\cdot)$ can be constructed similarly by replacing Q and ΔA by P and ΔA^T . Hence these bounds will not be discussed separately. For convenience in discussing the set \mathcal{U} , we shall use the terms *rectangle* and *ellipse* to refer to closed regions bounded by such figures in multiple dimensions. As usual, a polytope is the convex hull of a finite number of points.

5.1 Type I Bounds

We begin by constructing bounds $\Omega(\cdot)$ which exploit only the symmetry of the Lyapunov terms $\Delta A Q + Q \Delta A^T$. First we require the following definition of a function of a symmetric matrix as an extension of a real-valued function ([40], p. 300). Specifically, if $f: \mathbb{R} \rightarrow \mathbb{R}$ then (with a minor abuse of notation) $f: \mathbb{S}^n \rightarrow \mathbb{S}^n$ can be defined by setting

$$f(S) \triangleq U f(D) U^T,$$

where $S = U D U^T$, U is orthogonal, D is real diagonal, and $f(D)$ is the diagonal matrix obtained by applying f to each diagonal element of D . Note that if f is the polynomial $f(x) = \sum_{i=0}^r a_i x^i$ then $f(S) = \sum_{i=0}^r a_i S^i$. Note also that if $f(x) = |x|$ then $f(S) = (S^2)^{\frac{1}{2}}$, where $(\cdot)^{\frac{1}{2}}$ denotes the (unique) nonnegative-definite square root. As in [41], p. 262, we use the notation $|S|$ to denote $(S^2)^{\frac{1}{2}}$. Finally, note that if $f: \mathbb{R} \rightarrow \mathbb{R}$ and $g: \mathbb{R} \rightarrow \mathbb{R}$ are such that $f(x) \leq g(x)$, $x \in \mathbb{R}$, then $f(S) \leq g(S)$, $S \in \mathbb{S}^n$.

As a concretization of the uncertainty set \mathcal{U} , consider the set

$$\mathcal{U}_1 \triangleq \{ \Delta A \in \mathbb{R}^{n \times n} : \Delta A = \sum_{i=1}^p \sigma_i A_i, \quad |\sigma_i| \leq \delta_i, \quad i = 1, \dots, p \}, \quad (5.1)$$

where, for $i = 1, \dots, p$: $A_i \in \mathbb{R}^{n \times n}$ is a given matrix denoting the structure of the parametric uncertainty; σ_i is a real uncertain parameter; and δ_i denotes the range of parameter uncertainty. Clearly, the multidimensional set of uncertain parameters $(\sigma_1, \dots, \sigma_p)$ is the rectangle $[-\delta_1, \delta_1] \times \dots \times [-\delta_p, \delta_p]$ and \mathcal{U}_1 is a symmetric polytope of matrices in $\mathbb{R}^{n \times n}$. Note that the symmetry of the uncertainty interval $[-\delta_i, \delta_i]$ entails no loss of generality since the nominal value of A can be redefined if necessary. Furthermore, it is also possible, without loss of generality, to define $\delta_i = 1$ by replacing A_i by $\delta_i A_i$. For clarity, however, we choose not to employ this scaling.

We begin by considering the absolute-value bound utilized by Chang and Peng in [9].

Proposition 5.1. The function

$$\Omega_1(Q) \triangleq \sum_{i=1}^p \delta_i |A_i Q + Q A_i^T| \quad (5.2)$$

bounds \mathcal{U}_1 .

Proof. For $i = 1, \dots, p$ and $|\sigma_i| \leq \delta_i$,

$$\sigma_i(A_i Q + Q A_i^T) \leq |\sigma_i(A_i Q + Q A_i^T)| = |\sigma_i| |A_i Q + Q A_i^T| \leq \delta_i |A_i Q + Q A_i^T|.$$

Thus, summing over i yields

$$\Delta A Q + Q \Delta A^T = \sum_{i=1}^p \sigma_i(A_i Q + Q A_i^T) \leq \sum_{i=1}^p \delta_i |A_i Q + Q A_i^T|,$$

which yields (3.1) with $\Omega(\cdot) = \Omega_1(\cdot)$ and $\mathcal{U} = \mathcal{U}_1$. \square

Remark 5.1. It is tempting to prove Proposition 5.1 by writing

$$\sum_{i=1}^p \sigma_i(A_i Q + Q A_i^T) \leq \left| \sum_{i=1}^p \sigma_i(A_i Q + Q A_i^T) \right| \leq \sum_{i=1}^p |\sigma_i(A_i Q + Q A_i^T)|.$$

However, counterexamples show that the inequality $|M_1 + M_2| \leq |M_1| + |M_2|$ is not generally true for arbitrary symmetric matrices M_1, M_2 .

Remark 5.2. Because of its simplicity it is tempting to conjecture that $\Omega_1(\cdot)$ is the best bound for $\Delta A Q + Q \Delta A^T$ over the set \mathcal{U}_1 . To show that this is not the case, let $Q = \frac{1}{2} I_2$, $p = 1$, $A_1 = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$, and $\delta_1 = 1$. Then $\sigma_1(A_1 Q + Q A_1^T) \leq \delta_1 |A_1 Q + Q A_1^T| = I_2$, $|\sigma_1| \leq 1$. However, it is also true that $\sigma_1(A_1 Q + Q A_1^T) \leq \begin{bmatrix} 2 & 3/2 \\ 3/2 & 2 \end{bmatrix}$, $|\sigma_1| \leq 1$. Neither bound, however, is an overbound for the other. This is a consequence of the fact that the nonnegative-definite matrix ordering is only a partial order.

As mentioned above, an overbound for $\Omega_1(\cdot)$ will also bound \mathcal{U}_1 .

Lemma 5.1. For $i = 1, \dots, p$, let $f_i : \mathbb{R} \rightarrow \mathbb{R}$ satisfy

$$f_i(x) \geq |x|, \quad x \in \mathbb{R}. \quad (5.3)$$

Then the function

$$\Omega_2(Q) \triangleq \sum_{i=1}^p \delta_i f_i(A_i Q + Q A_i^T) \quad (5.4)$$

is an overbound for $\Omega_1(\cdot)$ and hence also bounds \mathcal{U}_1 .

One particular choice of f_i satisfying (5.3) will be considered here, namely, the polynomial

$$f_i(x) = \frac{1}{4}\beta_i + \beta_i^{-1}x^2, \quad (5.5)$$

where β_i is an arbitrary positive constant. Thus $\Omega_2(\cdot)$ has the following specialization.

Corollary 5.1. Let β_1, \dots, β_p be arbitrary positive constants. Then the function

$$\Omega_3(Q) \triangleq \frac{1}{4} \sum_{i=1}^p \delta_i \beta_i I_n + \sum_{i=1}^p (\delta_i / \beta_i) (A_i Q + Q A_i^T)^2 \quad (5.6)$$

is an overbound for $\Omega_1(\cdot)$ and hence also bounds \mathcal{U}_1 .

Although overbounding $\Omega_1(\cdot)$ by $\Omega_3(\cdot)$ results in a looser bound for \mathcal{U}_1 , it turns out that $\Omega_3(\cdot)$ actually bounds a set which is larger than \mathcal{U}_1 . Specifically, in place of \mathcal{U}_1 consider

$$\mathcal{U}_2 \triangleq \{\Delta A \in \mathbb{R}^{n \times n} : \Delta A = \sum_{i=1}^p \sigma_i A_i, \sum_{i=1}^p \sigma_i^2 / \alpha_i^2 \leq 1\}, \quad (5.7)$$

where $\alpha_1, \dots, \alpha_p$ are given positive constants. Note that (5.7) replaces the rectangle of uncertain parameters $(\sigma_1, \dots, \sigma_p)$ by an ellipse. Thus the set \mathcal{U}_2 of matrix perturbations is an ellipse of matrices in $\mathbb{R}^{n \times n}$ in contrast to the polytope \mathcal{U}_1 . Of course, $\mathcal{U}_1 = \mathcal{U}_2$ if $p = 1$ and $\alpha_1 = \delta_1$. Again it is possible to take $\alpha_i = 1$ without loss of generality by replacing A_i by $\alpha_i A_i$. We again choose not to do this, however. The following lemma provides a convenient characterization of the relationship between the rectangle \mathcal{U}_1 and the ellipse \mathcal{U}_2 .

Proposition 5.2. Suppose \mathcal{U}_1 is defined by the positive constants $\delta_1, \dots, \delta_p$ and let \mathcal{U}_2 be characterized by

$$\alpha_i = (\alpha \delta_i / \beta_i)^{\frac{1}{2}}, \quad i = 1, \dots, p, \quad (5.8)$$

where α is defined by

$$\alpha = \sum_{i=1}^p \delta_i \beta_i \quad (5.9)$$

and β_1, \dots, β_p are arbitrary positive constants. Then \mathcal{U}_2 contains \mathcal{U}_1 . Specifically, the ellipse \mathcal{U}_2 circumscribes the polytope \mathcal{U}_1 . Furthermore, $\Omega_3(\cdot)$ actually bounds \mathcal{U}_2 .

Proof. If $|\sigma_i| \leq \delta_i$, $i = 1, \dots, p$, then it follows from (5.8) that

$$\sum_{i=1}^p \sigma_i^2 / \alpha_i^2 = \alpha^{-1} \sum_{i=1}^p \beta_i \sigma_i^2 / \delta_i \leq \alpha^{-1} \sum_{i=1}^p \beta_i \delta_i = 1.$$

If, in addition, $|\sigma_i| = \delta_i$, $i = 1, \dots, p$, then $\sum_{i=1}^p \delta_i^2 / \alpha_i^2 = 1$, which corresponds to a point on the boundary of the ellipse. To show that $\Omega_3(\cdot)$ actually bounds \mathcal{U}_2 note that

$$\begin{aligned} 0 &\leq \sum_{i=1}^p \left[\frac{1}{2} (\alpha_i^{\frac{1}{2}} \sigma_i / \alpha_i) I_n - (\alpha_i / \alpha_i^{\frac{1}{2}}) (A_i Q + Q A_i^T) \right]^2 \\ &= \frac{\alpha}{4} \sum_{i=1}^p (\sigma_i^2 / \alpha_i^2) I_n + \alpha^{-1} \sum_{i=1}^p \alpha_i^2 (A_i Q + Q A_i^T)^2 - (\Delta A Q + Q \Delta A^T). \end{aligned}$$

Since $\sum_{i=1}^p \sigma_i^2 / \alpha_i^2 \leq 1$ in \mathcal{U}_2 , it follows that

$$\Delta A Q + Q \Delta A^T \leq \frac{\alpha}{4} I_n + \alpha^{-1} \sum_{i=1}^p \alpha_i^2 (A_i Q + Q A_i^T)^2.$$

Utilizing (5.8) and (5.9) to substitute for α and α_i yields (3.1) with $\Omega(\cdot) = \Omega_3(\cdot)$ and $\mathcal{U} = \mathcal{U}_2$. \square

Proposition 5.2 shows that each choice of constants $\beta_1, \dots, \beta_p > 0$ leads to a particular ellipse \mathcal{U}_2 which contains the polytope \mathcal{U}_1 . Furthermore, $\Omega_3(\cdot)$, which by Corollary 5.1 bounds \mathcal{U}_1 , actually bounds the larger set \mathcal{U}_2 . For convenience, we now dispense with the constants β_1, \dots, β_p which relate the rectangle \mathcal{U}_1 to the ellipse \mathcal{U}_2 and we characterize $\Omega_3(\cdot)$ entirely in terms of $\alpha_1, \dots, \alpha_p$.

Corollary 5.2. Let α be an arbitrary positive constant. Then the function

$$\Omega_4(Q) \triangleq \frac{\alpha}{4} I_n + \alpha^{-1} \sum_{i=1}^p \alpha_i^2 (A_i Q + Q A_i^T)^2 \quad (5.10)$$

bounds \mathcal{U}_2 .

Remark 5.3. Within the context of Corollary 5.2, the positive constant α plays no role in defining the set \mathcal{U}_2 , although $\Omega_4(\cdot)$ is guaranteed to bound \mathcal{U}_2 for all choices of α . It can be expected, however, that certain choices of α provide better bounds than other choices. This will be seen by example in Section 10.

5.2 Type II Bounds

We now consider additional bounds for \mathcal{U} which exploit the structure of Q . For these bounds the natural uncertainty set is given by \mathcal{U}_2 .

Proposition 5.3. Let α be an arbitrary positive number and, for each $Q \in \mathbb{R}^n$, let $Q_1 \in \mathbb{R}^{n \times m}$ and $Q_2 \in \mathbb{R}^{m \times n}$ satisfy

$$Q = Q_1 Q_2. \quad (5.11)$$

Then the function

$$\Omega_5(Q) \triangleq \alpha Q_2^T Q_2 + \alpha^{-1} \sum_{i=1}^p \alpha_i^2 A_i Q_1 Q_1^T A_i^T \quad (5.12)$$

bounds \mathcal{U}_2 .

Proof. Note that

$$\begin{aligned} 0 &\leq \sum_{i=1}^p [(\alpha_i^{1/2} \sigma_i / \alpha_i) Q_2^T - (\alpha_i / \alpha_i^{1/2}) A_i Q_1] [(\alpha_i^{1/2} \sigma_i / \alpha_i) Q_2^T - (\alpha_i / \alpha_i^{1/2}) A_i Q_1]^T \\ &= \alpha \sum_{i=1}^p (\sigma_i^2 / \alpha_i^2) Q_2^T Q_2 + \alpha^{-1} \sum_{i=1}^p \alpha_i^2 A_i Q_1 Q_1^T A_i^T - \sum_{i=1}^p \sigma_i (A_i Q + Q A_i^T), \end{aligned}$$

which, since $\sum_{i=1}^p \sigma_i^2 / \alpha_i^2 \leq 1$, yields (3.1) with $\Omega(\cdot) = \Omega_5(\cdot)$ and $\mathcal{U} = \mathcal{U}_2$. \square

We consider three specializations of $\Omega_5(\cdot)$. Specifically, we set $m = n$ and define

$$Q_1 = Q, \quad Q_2 = I_n, \quad (5.13)$$

$$Q_1 = Q_2 = Q^{1/2}, \quad (5.14)$$

$$Q_1 = I_n, \quad Q_2 = Q. \quad (5.15)$$

Corollary 5.3. Let α be an arbitrary positive number. Then the functions

$$\Omega_6(Q) \triangleq \alpha I_n + \alpha^{-1} \sum_{i=1}^p \alpha_i^2 A_i Q^2 A_i^T, \quad (5.16)$$

$$\Omega_7(Q) \triangleq \alpha Q + \alpha^{-1} \sum_{i=1}^p \alpha_i^2 A_i Q A_i^T, \quad (5.17)$$

$$\Omega_8(Q) \triangleq \alpha Q^2 + \alpha^{-1} \sum_{i=1}^p \alpha_i^2 A_i A_i^T \quad (5.18)$$

bound \mathcal{U}_2 .

Remark 5.3. Note that the term $A_i Q^2 A_i^T$ appearing in $\Omega_6(\cdot)$ also appears in $\Omega_4(\cdot)$. Furthermore, both $\Omega_4(\cdot)$ and $\Omega_6(\cdot)$ involve a term proportional to I_n . Despite these similarities, neither bound $\Omega_4(\cdot)$ nor $\Omega_6(\cdot)$ is an overbound for the other.

Remark 5.4. The bound $\Omega_7(\cdot)$ given by (5.17) has the distinction that it is linear in Q . This bound was originally studied in [23,27] for systems with multiplicative white noise and was shown to yield robust stability and performance in [33,35]. A similar bound was studied in [34].

Remark 5.5. Using (5.11) additional bounds can be developed. For example, by setting

$$Q_1 = Q^{\frac{1}{2}}, \quad Q_2 = Q^{\frac{1}{2}}, \quad (5.19)$$

$\Omega_8(Q)$ becomes

$$\Omega_9(Q) = \alpha Q^{\frac{1}{2}} + \alpha^{-1} \sum_{i=1}^p \alpha_i^2 A_i Q^{\frac{1}{2}} A_i^T. \quad (5.20)$$

Remark 5.6. When $p = 1$ and α is replaced by $\alpha\alpha_i$, $\Omega_7(\cdot)$ becomes

$$\Omega_7'(Q) = \alpha_1[\alpha Q + \alpha^{-1} A_1 Q A_1^T].$$

Utilizing a sum of such terms with $\alpha_i = \delta_i$ can be used to bound the smaller rectangular set \mathcal{U}_1 . Similar remarks apply to $\Omega_6(\cdot)$, $\Omega_8(\cdot)$, and $\Omega_9(\cdot)$.

5.3 Type III Bounds

We now consider bounds which exploit the structure of ΔA itself. It turns out that these bounds permit consideration of an uncertainty set \mathcal{U} which is larger than \mathcal{U}_2 . Specifically, define

$$\mathcal{U}_3 \triangleq \{\Delta A \in \mathbb{R}^{n \times n} : \Delta A = MN, \quad MM^T \leq D, \quad N^T N \leq E\}, \quad (5.21)$$

where $M \in \mathbb{R}^{n \times r}$ and $N \in \mathbb{R}^{r \times n}$ are uncertain matrices, r is an arbitrary positive integer, and $D, E \in \mathbb{R}^{n \times n}$ are given uncertainty bounds. The bound $\Omega_{10}(\cdot)$ for \mathcal{U}_3 is given by the following result.

Proposition 5.4. Let α be an arbitrary positive constant. Then the function

$$\Omega_{10}(Q) \triangleq \alpha^{-1} D + \alpha Q E Q \quad (5.22)$$

bounds \mathcal{U}_3 .

Proof. Note that

$$\begin{aligned} 0 &\leq [\alpha^{-\frac{1}{2}} M - \alpha^{\frac{1}{2}} Q N^T][\alpha^{-\frac{1}{2}} M - \alpha^{\frac{1}{2}} Q N^T]^T \\ &= \alpha^{-1} M M^T + \alpha Q N^T N Q - [M N Q + Q (M N)^T] \\ &\leq \alpha^{-1} D + \alpha Q E Q - (\Delta A Q + Q \Delta A^T), \end{aligned}$$

which yields (3.1) with $\Omega(\cdot) = \Omega_{10}(\cdot)$ and $\mathcal{U} = \mathcal{U}_3$. \square

Remark 5.7. The bound $\Omega_{10}(\cdot)$ was developed independently in [29] for robust analysis and in [25,28] for robust full-state feedback. Applications to fixed-order dynamic compensation are given in [36].

Remark 5.8. Without loss of generality we can set $\alpha = 1$ in (5.22) by replacing D and E by $\alpha^{-1}D$ and αE , respectively. Again for clarity we have chosen not to employ this normalization.

Note that $\Omega_{10}(\cdot)$ is an overbound for $\Omega_8(\cdot)$ when D and E are chosen to satisfy

$$\sum_{i=1}^p \alpha_i^2 A_i A_i^T \leq D, \quad I_n \leq E. \quad (5.23)$$

Hence, in this case $\Omega_{10}(\cdot)$ necessarily bounds \mathcal{U}_2 . As in the case of $\Omega_3(\cdot)$ overbounding $\Omega_1(\cdot)$, we should not be surprised to find that $\Omega_{10}(\cdot)$ actually bounds a set larger than \mathcal{U}_2 . Indeed, we now show that \mathcal{U}_2 is actually a very special subset of \mathcal{U}_3 when D and E defining \mathcal{U}_2 satisfy (5.23).

Proposition 5.5. Suppose D and E satisfy (5.23). Then \mathcal{U}_2 is a subset of \mathcal{U}_3 and thus the bound $\Omega_{10}(\cdot)$ for \mathcal{U}_3 also bounds \mathcal{U}_2 .

Proof. If $\Delta A \in \mathcal{U}_2$ then $\Delta A = \sum_{i=1}^p \sigma_i A_i$, where $\sum_{i=1}^p \sigma_i^2 / \alpha_i^2 \leq 1$. Alternatively, we can write $\Delta A = MN$, where $r = pn$ and

$$M = [\alpha_1 A_1 \dots \alpha_p A_p], \quad N = \begin{bmatrix} (\sigma_1 / \alpha_1) I_n \\ \vdots \\ (\sigma_p / \alpha_p) I_n \end{bmatrix}. \quad (5.24)$$

Note that with D and E satisfying (5.23) and M and N defined by (5.24) it follows that $MM^T \leq D$ and $N^T N \leq E$. Thus $\Delta A \in \mathcal{U}_3$. \square

The following result provides general conditions under which $\Omega_{10}(\cdot)$ bounds \mathcal{U}_2 .

Proposition 5.6. Suppose $A_i = D_i E_i$, $i = 1, \dots, p$, where $D_i \in \mathbb{R}^{n \times n_i}$ and $E_i \in \mathbb{R}^{n_i \times n}$, and suppose that

$$\sum_{i=1}^p \alpha_i^2 D_i D_i^T \leq D, \quad \sum_{i=1}^p E_i^T E_i \leq E. \quad (5.25)$$

Then \mathcal{U}_2 is a subset of \mathcal{U}_3 and thus the bound $\Omega_{10}(\cdot)$ for \mathcal{U}_3 also bounds \mathcal{U}_2 .

Proof. The result follows as in the proof Proposition 5.5. \square

Remark 5.9. When $p = 1$, $D = \alpha_1^2 D_1 D_1^T$ and $E = E_1^T E_1$, it is convenient to replace α by $\alpha \alpha_1$ so that $\Omega_{10}(\cdot)$ becomes

$$\Omega'_{10}(Q) = \alpha_1 [\alpha^{-1} D_1 D_1^T + \alpha Q E_1^T E_1 Q]. \quad (5.26)$$

In certain situations it is desirable to consider subsets of \mathcal{U}_3 of special structure. For example, define

$$\mathcal{U}_4 \triangleq \{\Delta A \in \mathbb{R}^{n \times n} : \Delta A = D_0 M N E_0, \quad D_0 M M^T D_0^T \leq D, \quad E_0^T N^T N E_0 \leq E\},$$

where $D_0 \in \mathbb{R}^{n \times n_1}$ and $E_0 \in \mathbb{R}^{n_2 \times n}$ are known matrices denoting the structure of the uncertainty, and $M \in \mathbb{R}^{n_1 \times r}$ and $N \in \mathbb{R}^{r \times n_2}$ are uncertain matrices. Finer structure can be included within \mathcal{U}_4 by replacing $D_0 M N E_0$ by a sum of terms $D_i M_i N_i E_i$, where D_i, E_i are known and M_i, N_i are uncertain ([29]). Note, however, that even though \mathcal{U}_4 is a proper subset of \mathcal{U}_3 , the *form* of the bound $\Omega_{10}(\cdot)$ does not change. Thus such refinements render the bound $\Omega_{10}(\cdot)$ conservative with respect to \mathcal{U}_4 since the *larger* uncertainty set \mathcal{U}_3 is actually being bounded.

6. Robust Stability and Performance via Modified Lyapunov Equations

We now combine the principal results of Sections 3, 4 and 5 to obtain a series of conditions guaranteeing robust stability and performance. In particular, we focus on bounds $\Omega_1, \Omega_4, \Omega_6, \Omega_7$, and Ω_{10} . For simplicity we shall frequently assume that V is positive definite so that (3.3) is satisfied. In this case it follows that the solution Q of (3.2) is positive definite. Our first result is a corollary of Theorem 3.1 with $\Omega(\cdot) = \Omega_1(\cdot)$ and $\mathcal{U} = \mathcal{U}_1$.

Theorem 6.1. Let $V \in \mathbb{P}^n$, $\delta_1, \dots, \delta_p > 0$, and suppose there exists $Q \in \mathbb{P}^n$ satisfying

$$0 = A Q + Q A^T + \sum_{i=1}^p \delta_i |A_i Q + Q A_i^T| + V. \quad (\text{MLE1})$$

Then $A + \Delta A$ is asymptotically stable for all $\Delta A \in \mathcal{U}_1$, and

$$J_S(\mathcal{U}_1) \leq \text{tr } Q R, \quad (6.1)$$

$$J_D(\mathcal{U}_1) \leq \lambda_{\max}(Q R). \quad (6.2)$$

For the next result define

$$A_\alpha \triangleq A + \frac{\alpha}{2} I_n \quad (6.3)$$

and

$$\gamma_i \triangleq \alpha_i^2 / \alpha, \quad i = 1, \dots, p. \quad (6.4)$$

Setting $\Omega(\cdot) = \Omega_4(\cdot), \Omega_6(\cdot), \Omega_7(\cdot)$ and $\mathcal{U} = \mathcal{U}_2$ yields the following corollary of Theorem 3.1.

Theorem 6.2. Let $V \in \mathbb{P}^n$, $\alpha, \alpha_1, \dots, \alpha_p > 0$, and suppose there exists $Q \in \mathbb{P}^n$ satisfying either

$$0 = A Q + Q A^T + \sum_{i=1}^p \gamma_i (A_i Q + Q A_i^T)^2 + \frac{\alpha}{4} I_n + V, \quad (\text{MLE2})$$

$$0 = AQ + QA^T + \sum_{i=1}^p \gamma_i A_i Q^2 A_i^T + \alpha I_n + V, \quad (\text{MLE3})$$

or

$$0 = A_\alpha Q + QA_\alpha^T + \sum_{i=1}^p \gamma_i A_i Q A_i^T + V. \quad (\text{MLE4})$$

Then $A + \Delta A$ is asymptotically stable for all $\Delta A \in \mathcal{U}_2$, and

$$J_S(\mathcal{U}_2) \leq \text{tr } QR, \quad (6.5)$$

$$J_D(\mathcal{U}_2) \leq \lambda_{\max}(QR). \quad (6.6)$$

Next we set $\Omega(\cdot) = \Omega_{10}(\cdot)$ and $\mathcal{U} = \mathcal{U}_3$.

Theorem 6.3. Let $V \in \mathbb{P}^n$, $D \in \mathbb{N}^n$, and $E \in \mathbb{N}^n$, and suppose there exists $Q \in \mathbb{P}^n$ satisfying

$$0 = AQ + QA^T + \alpha QEQ + \alpha^{-1}D + V. \quad (\text{MLE5})$$

Then $A + \Delta A$ is asymptotically stable for all $\Delta A \in \mathcal{U}_3$, and

$$J_S(\mathcal{U}_3) \leq \text{tr } QR, \quad (6.7)$$

$$J_D(\mathcal{U}_3) \leq \lambda_{\max}(QR). \quad (6.8)$$

Additional sufficient conditions can be obtained by considering "mixed" bounds. That is, one can construct modified Lyapunov equations by combining two or more different bounds. Although mixed bounds will not be considered further in this paper, we present one such result for illustrative purposes.

Theorem 6.4. Let $V \in \mathbb{P}^n$, $\delta_1, \dots, \delta_p > 0$, $D \in \mathbb{N}^n$, and $E \in \mathbb{N}^n$, and suppose there exists $Q \in \mathbb{P}^n$ satisfying

$$0 = AQ + QA^T + \sum_{i=1}^p \delta_i |A_i Q + Q A_i^T| + \alpha QEQ + \alpha^{-1}D + V. \quad (\text{MLE1,5})$$

Then $A + \Delta A$ is asymptotically stable for all $\Delta A \in \mathcal{U}_1 + \mathcal{U}_3$, and

$$J_S(\mathcal{U}_1 + \mathcal{U}_3) \leq \text{tr } QR, \quad (6.9)$$

$$J_D(\mathcal{U}_1 + \mathcal{U}_3) \leq \lambda_{\max}(QR). \quad (6.10)$$

As noted previously, the bound $\Lambda(\cdot)$ can readily be constructed by replacing ΔA by ΔA^T in the definitions of $\Omega_1(\cdot)$ through $\Omega_{10}(\cdot)$. Denote these bounds by $\Lambda_1(\cdot)$ through $\Lambda_{10}(\cdot)$, respectively. For illustration we state the dual of Theorem 6.1 involving $\Lambda_1(\cdot)$. The dual versions of MLE1 through MLE5 will be denoted by MLED1 through MLED5.

Theorem 6.5. Let $R \in \mathbb{P}^n$, $\delta_1, \dots, \delta_p > 0$ and suppose there exists $P \in \mathbb{P}^n$ satisfying

$$0 = A^T P + P A + \sum_{i=1}^p \delta_i |A_i^T P + P A_i| + R. \quad (\text{MLED1})$$

Then $A + \Delta A$ is asymptotically stable for all $\Delta A \in \mathcal{U}_1$, and

$$J_S(\mathcal{U}_1) \leq \text{tr } PV, \quad (6.11)$$

$$\hat{J}_D(\mathcal{U}_1) \leq \lambda_{\max}(PV). \quad (6.12)$$

It is reasonable to expect that the sufficient conditions given by Theorems 3.1 and 4.1 are generally different. For example, the modified Lyapunov equations and their duals need not both possess a solution, while the bounds $\text{tr } QR$ and $\text{tr } PV$ need not be equal. An exception is the case in which $\Omega(\cdot) = \Omega_7(\cdot)$ and $\Lambda(\cdot) = \Lambda_7(\cdot)$. Note that the dual of MLE4 is given by

$$0 = A_\alpha^T P + P A_\alpha + \sum_{i=1}^p \gamma_i A_i^T P A_i + V. \quad (\text{MLED4})$$

Proposition 6.1. Let $\alpha_1, \dots, \alpha_p > 0$ and assume there exist $Q, P \in \mathbb{N}^n$ satisfying MLE4 and MLED4, respectively. Then

$$\text{tr } QR = \text{tr } PV. \quad (6.13)$$

Proof. Note that

$$\begin{aligned} \text{tr } QR &= -\text{tr } Q(A_\alpha^T P + P A_\alpha + \sum_{i=1}^p \gamma_i A_i^T P A_i) \\ &= -\text{tr } P(A_\alpha Q + Q A_\alpha^T + \sum_{i=1}^p \gamma_i A_i Q A_i^T) \\ &= \text{tr } PV. \quad \square \end{aligned}$$

Remark 6.1. By setting $\Omega(\cdot) = \Omega_7(\cdot)$ and $\Lambda(\cdot) = \Lambda_7(\cdot)$ it follows from (4.14) that

$$\text{tr } Q_0(\alpha P + \sum_{i=1}^p \gamma_i A_i^T P A_i) = \text{tr } P_0(\alpha Q + \sum_{i=1}^p \gamma_i A_i Q A_i^T). \quad (6.14)$$

7. Existence, Uniqueness, and Monotonicity of Solutions to the Modified Lyapunov Equations

It is important to stress that the sufficient conditions for robustness given by Theorems 6.1–6.5 assume only that there exist nonnegative-definite solutions Q, P satisfying the modified Lyapunov equations. Indeed, no *explicit* assumptions on the problem data A, V, R , and \mathcal{U} were utilized for assuring robust stability and performance. In applying Theorems 6.1–6.5 to specific problems it thus suffices to show that a nonnegative-definite solution Q *exists* in order to obtain robust stability, while, for robust performance, the bounds (6.1), (6.2), (6.5)–(6.8) require explicit knowledge of Q . Thus, any computational method which yields a nonnegative-definite solution will suffice to guarantee both robust stability and performance.

Before considering the numerical solution of the modified Lyapunov equations, several relevant issues require discussion. For example, before seeking to compute solutions to MLE1–MLE5 it would be desirable to determine a priori whether these equations actually possess nonnegative-definite solutions. For example, it may be useful to obtain sufficient and/or necessary conditions for the *existence* of nonnegative-definite solutions. Thus, if the sufficient conditions are satisfied then existence (and hence robustness) is assured, while if the necessary conditions are *not* satisfied then existence is ruled out. If, on the other hand, either the sufficient conditions are not satisfied or the necessary conditions *are* satisfied, then nothing can be surmised. Finally, such conditions need to be easily verifiable and reasonably nonconservative since otherwise it would be more prudent to attempt to numerically solve the modified Lyapunov equation itself.

It is quite possible that at least some of the modified Lyapunov equations possess multiple nonnegative-definite solutions. In this case one may seek the minimal solution (i.e., the smallest with respect to the nonnegative-definite matrix ordering) in order to minimize the performance bounds. If multiple solutions exist, none of which is minimal, then the best bound would depend upon the matrix R .

Since the matrix Q determines the performance bound, it is reasonable to expect that Q is *monotonic* in \mathcal{U} . That is, if \mathcal{U} decreases in size, then the solution Q is more likely to exist while decreasing in the nonnegative-definite matrix ordering. For example, consider \mathcal{U}'_1 characterized by δ'_i , where $\delta'_i \leq \delta_i$, $i = 1, \dots, p$. Then one might expect $Q' \leq Q$ where Q' is the solution to MLE1 with δ_i replaced by δ'_i . Finally, monotonicity with respect to V should also be expected.

Because of linearity, the analysis of bound $\Omega_7(\cdot)$ is simplest and it is possible to obtain necessary and sufficient conditions for the existence of solutions to MLE4. The basic tool required is the Kronecker matrix algebra ([42]). For convenience, define

$$A \triangleq A_\alpha \oplus A_\alpha + \sum_{i=1}^p \gamma_i A_i \otimes A_i. \quad (7.1)$$

Proposition 7.1. If $V \in \mathbb{R}^n$ and A is asymptotically stable, then there exists a unique $n \times n$ Q satisfying MLE4, and $Q \geq 0$. Conversely, if for all $V \in \mathbb{R}^n$ there exists $Q \geq 0$ satisfying MLE4, then A is asymptotically stable.

Proof. Since MLE4 is equivalent to

$$Q = -\text{vec}^{-1}[A^{-1}\text{vec } V], \quad (7.2)$$

existence and uniqueness hold. Here, vec and vec^{-1} denote the column-stacking operation ([42]) and its inverse. To prove that Q is nonnegative definite, we rewrite (7.2) as

$$Q = \int_0^\infty \text{vec}^{-1}[e^{At}\text{vec } V]dt \quad (7.3)$$

and show that the integrand is nonnegative-definite for all $t \in [0, \infty)$. [Note that the following argument does not require that A be stable.] Using the exponential product formula, the exponential in (7.3) can be written as

$$e^{At} = \lim_{k \rightarrow \infty} \left\{ \exp \left[\frac{1}{k} (A_\alpha \oplus A_\alpha) t \right] \exp \left[\frac{1}{k} \sum_{i=1}^p \gamma_i (A_i \otimes A_i) t \right] \right\}^k. \quad (7.4)$$

For convenience, let S and N be $r \times r$ matrices with $N \geq 0$. Since (see [42])

$$\text{vec}^{-1}[(S \otimes S)\text{vec } N] = SNS^T \geq 0 \quad (7.5)$$

and

$$(S^k \otimes S^k)(S \otimes S) = S^{k+1} \otimes S^{k+1}, \quad (7.6)$$

it follows that

$$\text{vec}^{-1}[e^{S \otimes S}\text{vec } N] = \sum_{k=0}^{\infty} (k!)^{-1} S^k N S^{kT} \geq 0. \quad (7.7)$$

Furthermore,

$$\text{vec}^{-1}[e^{S \otimes S}\text{vec } N] = \text{vec}^{-1}[(e^S \otimes e^S)\text{vec } N] = e^S N e^{S^T} \geq 0. \quad (7.8)$$

Applying (7.7) and (7.8) alternately with (7.4) and using induction on k it follows that the integrand of (7.3) is nonnegative definite. To prove the converse, note that it follows from MLE4 that Q satisfies

$$Q = \text{vec}^{-1}[e^{A^t} \text{vec } Q] + \int_0^t \text{vec}^{-1}[e^{A^s} \text{vec } V] ds, \quad t \in [0, \infty). \quad (7.9)$$

Since the integral term on the right hand side of (7.9) is nonnegative definite, is bounded from above by Q , and $V \in \mathbb{R}^n$ is arbitrary, it follows that A is asymptotically stable. \square

We now show that if A is asymptotically stable then actually A_α is asymptotically stable. This shows that the assumption that A is asymptotically stable is consistent with the original hypothesis that A is asymptotically stable.

Proposition 7.2. Assume A is asymptotically stable, let $\alpha'_i \in [0, \alpha_i]$, $i = 1, \dots, p$, and define

$$A' \triangleq A_\alpha \oplus A_\alpha + \sum_{i=1}^p (\alpha_i'^2 / \alpha) A_i \otimes A_i.$$

Then A' is also asymptotically stable. In particular, A_α and A are asymptotically stable.

Proof. Let $V \in \mathbb{R}^n$ and let Q be the unique, nonnegative-definite solution of MLE4. Equivalently, Q satisfies

$$0 = A_\alpha Q + Q A_\alpha^T + \sum_{i=1}^p (\alpha_i'^2 / \alpha) A_i Q A_i^T + V',$$

where

$$V' \triangleq \sum_{i=1}^p \alpha^{-1} (\alpha_i^2 - \alpha_i'^2) A_i Q A_i^T + V.$$

Since $V' \in \mathbb{R}^n$, the stability of A' now follows from the converse of Proposition 7.1. Finally, if $V \in \mathbb{R}^n$ then $\sum_{i=1}^p (\alpha_i'^2 / \alpha) A_i Q A_i^T + V'$ is positive definite and it follows from Lemma 12.2 of [39] that A_α , and hence A , is asymptotically stable. \square

Hence it follows from Proposition 7.2 that a necessary condition for A to be asymptotically stable is that

$$\alpha < 2 \max_{i=1, \dots, n} \text{Re } \lambda_i(A). \quad (7.10)$$

We now have the following monotonicity result.

Proposition 7.3. Let $\mathcal{U}'_2 \subset \mathcal{U}_2$, where \mathcal{U}'_2 is defined as in (5.7) with α_i replaced by $\alpha'_i \in [0, \alpha_i]$, $i = 1, \dots, p$. Furthermore, let $V \in \mathbb{R}^n$, assume A is asymptotically stable, and let $Q \in \mathbb{R}^n$

satisfy MLE4. Then there exists $Q' \in \mathbb{N}^n$ satisfying

$$0 = A_\alpha Q' + Q' A_\alpha^T + \sum_{i=1}^p (\alpha_i^2 / \alpha) A_i Q' A_i^T + V, \quad (7.11)$$

and, furthermore,

$$Q' \leq Q. \quad (7.12)$$

Consequently,

$$\text{tr } Q' R \leq \text{tr } Q R, \quad (7.13)$$

$$\lambda_{\max}(Q' R) \leq \lambda_{\max}(Q R). \quad (7.14)$$

Proof. Subtracting (7.11) from MLE4 yields

$$0 = A_\alpha(Q - Q') + (Q - Q') A_\alpha^T + \sum_{i=1}^p (\alpha_i^2 / \alpha) A_i (Q - Q') A_i^T + V',$$

where V' is defined in the proof of Proposition 7.2. Since, by the converse of Proposition 7.1, A' is asymptotically stable, $Q - Q' \geq 0$, which yields (7.12) and thus (7.13) and (7.14). \square

Returning now to the existence question, Proposition 7.1 shows that a solution to MLE4 exists so long as $\alpha_1, \dots, \alpha_p$ are sufficiently small that A remains stable for some $\alpha > 0$. To this end one can treat this as a stability perturbation problem and apply results from [3]. Within our modified Lyapunov equation approach we have the following related result. For this and the following result let $\|\cdot\|$ denote an arbitrary vector norm and the corresponding induced matrix norm.

Proposition 7.4. If

$$\|(A \oplus A)^{-1}(\alpha I_{n^2} + \alpha^{-1} \sum_{i=1}^p \alpha_i^2 A_i \otimes A_i)\| < 1, \quad (7.15)$$

then for all $V \in \mathbb{N}^n$ there exists $Q \in \mathbb{N}^n$ satisfying MLE4 and hence A is asymptotically stable.

Proof. Define $\{Q_k\}_{k=0}^\infty$ where Q_0 satisfies (3.14) and Q_{k+1} satisfies

$$0 = A Q_{k+1} + Q_{k+1} A^T + \Omega_7(Q_k) + V.$$

Note that $Q_k \geq 0$, $k = 1, 2, \dots$. Hence it follows that

$$\text{vec } Q_{k+1} - \text{vec } Q_k = -(A \oplus A)^{-1} [\text{vec } \Omega_7(Q_k) - \text{vec } \Omega_7(Q_{k-1})]$$

and thus

$$\|\text{vec } Q_{k+1} - \text{vec } Q_k\| \leq \|(A \oplus A)^{-1}(\alpha I_{n^2} + \alpha^{-1} \sum_{i=1}^p \alpha_i^2 A_i \otimes A_i)\| \|\text{vec } Q_k - \text{vec } Q_{k-1}\|.$$

Using (7.15) it follows that $Q \triangleq \lim_{k \rightarrow \infty} Q_k$ exists. Thus $Q \geq 0$ and satisfies MLE4. Finally, by the converse of Proposition 7.1, \mathcal{A} is asymptotically stable. \square

Since MLE5 is nonlinear a slightly different approach is required for existence. For the following result let $\kappa, \beta > 0$ satisfy

$$\|e^{At}\| \leq ke^{-\beta t}, \quad t \geq 0, \quad (7.16)$$

where $\|\cdot\|$ denotes an arbitrary submultiplicative matrix norm, and define $\rho \triangleq 2\beta/\kappa^2$.

Proposition 7.5. Suppose $V \in \mathbb{R}^n$ and

$$4\alpha\|E\|\|\alpha^{-1}D + V\| < \rho^2. \quad (7.17)$$

Then there exists $Q \in \mathbb{R}^n$ satisfying MLE5.

Proof. Consider the sequence $\{Q_k\}_{k=0}^\infty$ where Q_0 satisfies (3.14) and Q_{k+1} is given by

$$0 = AQ_{k+1} + Q_{k+1}A^T + \alpha Q_k E Q_k + \alpha^{-1}D + V.$$

Clearly, $Q_k \geq 0$, $k = 0, 1, \dots$. Next we have

$$Q_{k+1} = \int_0^\infty e^{At} [\alpha Q_k E Q_k + \alpha^{-1}D + V] e^{A^T t} dt \quad (7.18)$$

which yields

$$\|Q_{k+1}\| \leq \alpha \rho^{-1} \|E\| \|Q_k\|^2 + \rho^{-1} \|\alpha^{-1}D + V\|. \quad (7.19)$$

Similarly, we obtain

$$\|Q_0\| \leq \rho^{-1} \|\alpha^{-1}D + V\|.$$

Now suppose that

$$\|Q_k\| \leq 2\rho^{-1} \|\alpha^{-1}D + V\|.$$

Then (7.17) and (7.19) imply

$$\begin{aligned} \|Q_{k+1}\| &\leq \alpha \rho^{-1} \|E\| [2\rho^{-1} \|\alpha^{-1}D + V\|]^2 + \rho^{-1} \|\alpha^{-1}D + V\| \\ &< 2\rho^{-1} \|\alpha^{-1}D + V\|. \end{aligned}$$

Thus $\|Q_k\| \leq 2\rho^{-1}\|\alpha^{-1}D + V\|$, $k = 0, 1, \dots$. Next, (7.18) yields

$$\begin{aligned} Q_{k+1} - Q_k &= \alpha \int_0^\infty e^{At} [Q_k E Q_k - Q_{k-1} E Q_{k-1}] e^{A^T t} dt \\ &= \alpha \int_0^\infty e^{At} [Q_k E (Q_k - Q_{k-1}) + (Q_k - Q_{k-1}) E Q_{k-1}] e^{A^T t} dt \end{aligned}$$

and thus

$$\begin{aligned} \|Q_{k+1} - Q_k\| &\leq \alpha \rho^{-1} \|E\| (\|Q_k\| + \|Q_{k-1}\|) \|Q_k - Q_{k-1}\| \\ &\leq 4\alpha \rho^{-2} \|E\| \|\alpha^{-1}D + V\| \|Q_k - Q_{k-1}\| \\ &\leq \varepsilon \|Q_k - Q_{k+1}\|, \end{aligned}$$

where $\varepsilon \triangleq 4\alpha \rho^{-2} \|E\| \|\alpha^{-1}D + V\|$. Since by (7.16) $\varepsilon < 1$, $\lim_{k \rightarrow \infty} Q_k$ exists, is nonnegative definite, and satisfies MLE5. \square

8. Additional Upper Bounds via Recursive Substitution

In this section we obtain additional upper bounds for $J_S(\mathcal{U})$ and $J_D(\mathcal{U})$ by utilizing a recursive substitution technique. The main idea involves rewriting (2.7) as

$$Q_{\Delta A} = -\text{vec}^{-1}\{(A \oplus A)^{-1}(\Delta A \oplus \Delta A) \text{vec } Q_{\Delta A}\} + Q_0 \quad (8.1)$$

and substituting this expression into the terms $\Delta A Q_{\Delta A} + Q_{\Delta A} \Delta A^T$ appearing in (2.7). This technique yields an equation which is, as expected, equivalent to (2.7) but which permits the development of additional bounds. As will be seen, the ability to develop new bounds exploits the fact that the substitution technique leads to terms which are quadratic in ΔA . We begin the development with the following technical result which does not require the assumption that A is asymptotically stable.

Proposition 8.1. Suppose $A \oplus A$ is invertible and let $\Delta A \in \mathbb{R}^{n \times n}$. If $Q_{\Delta A}$ satisfies (2.7) then $Q_{\Delta A}$ also satisfies

$$\begin{aligned} 0 &= A Q_{\Delta A} + Q_{\Delta A} A^T - \text{vec}^{-1}[(\Delta A \oplus \Delta A)(A \oplus A)^{-1}(\Delta A \oplus \Delta A) \text{vec } Q_{\Delta A} \\ &\quad + (\Delta A \oplus \Delta A)(A \oplus A)^{-1} \text{vec } V] + V. \end{aligned} \quad (8.2)$$

Conversely, if $Q_{\Delta A}$ satisfies (8.2) and $(A - \Delta A) \oplus (A - \Delta A)$ is invertible, then $Q_{\Delta A}$ also satisfies (2.7).

Proof. To obtain (8.2) substitute (8.1) into (2.7) as noted above. Conversely, adding the zero term $(\Delta A \oplus \Delta A)(A \oplus A)^{-1}(A \oplus A) \text{vec } Q_{\Delta A} - (\Delta A \oplus \Delta A) \text{vec } Q_{\Delta A}$ to (8.2), it follows that (8.2) can be written as

$$0 = [(A - \Delta A) \oplus (A - \Delta A)](A \oplus A)^{-1}[(A + \Delta A) \oplus (A + \Delta A) \text{vec } Q_{\Delta A} + \text{vec } V],$$

which, under the invertibility assumption, implies that $Q_{\Delta A}$ satisfies (2.7). \square

The following result is analogous to Theorem 3.1.

Theorem 8.1. Suppose \mathcal{U} is symmetric, let $\Omega_0 \in \mathbb{R}^n$ satisfy

$$\Delta A Q_0 + Q_0 \Delta A^T \leq \Omega_0, \quad \Delta A \in \mathcal{U}, \quad (8.3)$$

where Q_0 satisfies (3.14), let $\hat{\Omega} : \mathbb{R}^n \rightarrow \mathbb{R}^n$ satisfy

$$-\text{vec}^{-1}[(\Delta A \oplus \Delta A)(A \oplus A)^{-1}(\Delta A \oplus \Delta A)\text{vec } Q] \leq \hat{\Omega}(Q), \quad \Delta A \in \mathcal{U}, \quad Q \in \mathbb{R}^n, \quad (8.4)$$

and suppose there exists $Q \in \mathbb{R}^n$ satisfying

$$0 = A Q + Q A^T + \hat{\Omega}(Q) + \Omega_0 + V. \quad (8.5)$$

Then

$$\left(A + \Delta A, \{ \hat{\Omega}(Q) + \text{vec}^{-1}[(\Delta A \oplus \Delta A)(A \oplus A)^{-1}(\Delta A \oplus \Delta A)\text{vec } Q] + \Omega_0 + V \}^{\frac{1}{2}} \right) \quad (8.6)$$

is stabilizable, $\Delta A \in \mathcal{U}$,

if and only if

$$A + \Delta A \text{ is asymptotically stable, } \Delta A \in \mathcal{U}. \quad (8.7)$$

In this case,

$$Q_{\Delta A} \leq Q, \quad \Delta A \in \mathcal{U}, \quad (8.8)$$

where $Q_{\Delta A}$ satisfies (2.7), and

$$J_S(\mathcal{U}) \leq \text{tr } QR, \quad (8.9)$$

$$J_D(\mathcal{U}) \leq \lambda_{\max}(QR). \quad (8.10)$$

Proof. The equivalence of (8.6) and (8.7) follows from (8.5) as in the proof of Theorem 3.1. Next (8.8) follows by comparing (8.5) and (8.2) while using (8.3) and (8.4). Since \mathcal{U} is assumed to be symmetric, it follows that $A - \Delta A$ is asymptotically stable, $\Delta A \in \mathcal{U}$, and hence $(A - \Delta A) \oplus (A - \Delta A)$ is invertible, $\Delta A \in \mathcal{U}$. Thus, the converse of Proposition 8.1 implies that $Q_{\Delta A}$ satisfying (8.2) also satisfies (2.7). Thus, the bound (8.8) can be used to obtain (8.9) and (8.10). \square

The principal difference between (8.4) and (3.1) is that ΔA appears linearly in (3.1) while it appears quadratically in (8.4). By exploiting this structure we can obtain new bounds for $Q_{\Delta A}$.

To simplify matters, we now consider the bound in (8.4) in two special cases. In the first case we set $\mathcal{U} = \mathcal{U}_1$ and $p = 1$ so that $\Delta A = \sigma_1 A_1$, $|\sigma_1| \leq \delta_1$. In this case (8.4) becomes

$$-\sigma_1^2 \text{vec}^{-1}[(A_1 \oplus A_1)(A \oplus A)^{-1}(A_1 \oplus A_1) \text{vec } Q] \leq \hat{\Omega}(Q), \quad |\sigma_1| \leq \delta_1, \quad Q \in \mathbb{N}^n. \quad (8.11)$$

One choice of $\hat{\Omega}(\cdot)$ which immediately suggests itself can be obtained by defining the matrix function $|\cdot|_+$ on the set of symmetric matrices by

$$|S|_+ \triangleq \frac{1}{2}(S + |S|), \quad (8.12)$$

which effectively replaces the negative eigenvalues of S by zeros. We shall thus utilize the fact that

$$\sigma_1^2 S \leq \delta_1^2 |S|_+, \quad |\sigma_1| \leq \delta_1, \quad (8.13)$$

for all symmetric S .

Corollary 8.1. Let $V \in \mathbb{P}^n$, $\mathcal{U} = \mathcal{U}_1$, $p = 1$, let $\Omega_0 \in \mathbb{N}^n$ satisfy (8.3), and suppose there exists $Q \in \mathbb{N}^n$ satisfying

$$0 = AQ + QA^T + \delta_1^2 | - \text{vec}^{-1}[(A_1 \oplus A_1)(A \oplus A)^{-1}(A_1 \oplus A_1) \text{vec } Q]|_+ + \Omega_0 + V. \quad (8.14)$$

Then (8.7)–(8.10) are satisfied.

For the next specialization we shall assume that

$$(\Delta A)A = A(\Delta A), \quad \Delta A \in \mathcal{U}, \quad (8.15)$$

which holds, for example, for modal systems with frequency uncertainty (see Section 10). It thus follows that $(A \oplus A)^{-1}(\Delta A \oplus \Delta A) = (\Delta A \oplus \Delta A)(A \oplus A)^{-1}$ and thus (8.4) can be rewritten as

$$\Delta A^2 \hat{Q} + 2\Delta A \hat{Q} \Delta A^T + \hat{Q} \Delta A^2 \leq \hat{\Omega}(Q), \quad \Delta A \in \mathbb{N}, \quad Q \in \mathbb{N}^n, \quad (8.16)$$

where $\hat{Q} \in \mathbb{N}^n$ satisfies

$$0 = A\hat{Q} + \hat{Q}A^T + Q. \quad (8.17)$$

Assuming in addition to (8.15) that $\Delta A = \sigma_1 A_1$, $|\sigma_1| \leq \delta_1$, (8.14) becomes

$$0 = AQ + QA^T + \delta_1^2 |A_1^2 \hat{Q} + 2A_1 \hat{Q} A_1^T + \hat{Q} A_1^2|_+ + \Omega_0 + V. \quad (8.18)$$

Remark 8.1. It is interesting to note that the left hand side of (8.16) is essentially of the same form as $\Omega_7(\cdot)$. Specifically, the term $\Delta A^2 \hat{Q} + \hat{Q} \Delta A^2 T$ is analogous to αQ while $2\Delta A \hat{Q} \Delta A^T$ is similar to $\alpha^{-1} \sum_{i=1}^p \alpha_i^2 A_i Q A_i^T$ appearing in $\Omega_7(\cdot)$. The form of the left hand side of (8.16) is also of interest since it is similar to terms which arise from a multiplicative white noise model with a Stratonovich correction. Specifically, while the term $\Delta A Q \Delta A^T$ arises from an Ito model ([33]), the new terms $\frac{1}{2} \Delta A^2$ can be viewed as a correction to the nominal A due to the Stratonovich interpretation of stochastic integration ([43]). These terms have interesting ramifications in designing robust controllers for flexible structures ([23]).

9. An Alternative Approach Yielding Upper and Lower Bounds

In this section we develop a variation on the results of Section 3 which has the additional benefit of yielding both upper and lower performance bounds. The basic approach was suggested by results obtained in [44]. To simplify the presentation we assume that if $\Delta A \in \mathcal{U}$ then $-\Delta A \in \mathcal{U}$. This symmetry assumption of course holds for all of the uncertainty sets considered in previous sections. The underlying idea involves bounding the deviation of $Q_{\Delta A}$ from Q_0 rather than bounding $Q_{\Delta A}$ directly.

Theorem 9.1. Let $\Omega_0 \in \mathbb{N}^n$ satisfy

$$\Delta A Q_0 + Q_0 \Delta A^T \leq \Omega_0, \quad \Delta A \in \mathcal{U}, \quad (9.1)$$

let $\Omega: \mathbb{N}^n \rightarrow \mathbb{N}^n$ be such that (3.1) is satisfied, and suppose there exists $\Delta Q \in \mathbb{N}^n$ satisfying

$$0 = A \Delta Q + \Delta Q A^T + \Omega(\Delta Q) + \Omega_0. \quad (9.2)$$

Then

$$(A + \Delta A, [\Omega_0 + \Omega(\Delta Q) - (\Delta A \Delta Q + \Delta Q \Delta A^T)]^{\frac{1}{2}}) \text{ is stabilizable, } \Delta A \in \mathcal{U}, \quad (9.3)$$

if and only if

$$A + \Delta A \text{ is asymptotically stable, } \Delta A \in \mathcal{U}. \quad (9.4)$$

In this case,

$$Q_0 - \Delta Q \leq Q_{\Delta A} \leq Q_0 + \Delta Q, \quad \Delta A \in \mathcal{U}, \quad (9.5)$$

where $Q_{\Delta A}$ is given by (2.7), and

$$\text{tr}(Q_0 - \Delta Q)R \leq J_s(\mathcal{U}) \leq \text{tr}(Q_0 + \Delta Q)R, \quad (9.6)$$

$$\lambda_{\max}[(Q_0 - \Delta Q)R] \leq J_D(\mathcal{U}) \leq \lambda_{\max}[(Q_0 + \Delta Q)R]. \quad (9.7)$$

Proof. Define

$$\Delta Q \triangleq Q_{\Delta A} - Q_0 \quad (9.8)$$

and subtract (3.14) from (2.7) to obtain

$$0 = (A + \Delta A)\Delta Q + \Delta Q(A + \Delta A)^T + \Delta A Q_0 + Q_0 \Delta A^T. \quad (9.9)$$

Now rewrite (9.2) as

$$0 = (A + \Delta A)\Delta Q + \Delta Q(A + \Delta A)^T + \Omega(\Delta Q) - (\Delta A \Delta Q + \Delta Q \Delta A^T) + \Omega_0. \quad (9.10)$$

Using (9.10), the equivalence of (9.3) and (9.4) is immediate as in the proof of Theorem 3.1. Next, subtracting (9.9) from (9.10) yields

$$\begin{aligned} 0 = (A + \Delta A)(\Delta Q - \Delta Q) + (\Delta Q - \Delta Q)(A + \Delta A)^T + \Omega(\Delta Q) - (\Delta A \Delta Q + \Delta Q \Delta A^T) \\ + \Omega_0 - (\Delta A Q_0 + Q_0 \Delta A^T). \end{aligned} \quad (9.11)$$

Using (3.1) and (9.1) it follows from (9.11) that

$$\Delta Q - \Delta Q \geq 0,$$

or, equivalently,

$$Q_{\Delta A} \leq Q_0 + \Delta Q. \quad (9.12)$$

To obtain the lower bound rewrite (9.9) as

$$0 = (A + \Delta A)(-\Delta Q) + (-\Delta Q)(A + \Delta A)^T - (\Delta A Q_0 + Q_0 \Delta A^T). \quad (9.13)$$

Note that because of the assumed symmetry of \mathcal{U} , (9.1) holds with ΔA appearing in the inequality replaced by $-\Delta A$. Hence it can be shown similarly that

$$\Delta Q + \Delta Q \geq 0,$$

or, equivalently,

$$Q_0 - \Delta Q \leq Q_{\Delta A}. \quad (9.14)$$

Finally, (9.6) follows immediately from (9.5) while (9.7) is a consequence of Theorem 4.3.1 of [40].

□

Remark 9.1. To compare the upper bound in (9.5) with (3.5), rewrite (9.2) as

$$0 = A(Q_0 + \Delta Q) + (Q_0 + \Delta Q)A^T + \Omega(\Delta Q) + \Omega_0 + V. \quad (9.15)$$

If $\Omega(\Delta Q) + \Omega_0 = \Omega(Q_0 + \Delta Q)$ then (9.15) has the same form as (3.2) and thus the two upper bounds are identical. This will be the case, for example, if $\Omega(\cdot) = \Omega_7(\cdot)$ and Ω_0 is chosen to be $\Omega_7(Q_0)$ since $\Omega_7(\cdot)$ is linear. If, for example, $\Omega(\Delta Q) + \Omega_0 < \Omega(Q_0 + \Delta Q)$ then the upper bound in (9.5) will be sharper. In any case it is clear that the individual treatment of ΔQ and Q_0 yields potentially new upper bounds.

Remark 9.2. Theorem 9.1 does not guarantee that the lower bound $Q_0 - \Delta Q$ for $Q_{\Delta A}$ is nonnegative definite. However, $Q_{\Delta A}$ is always nonnegative definite and thus the lower bound in (9.5) may be of limited usefulness. Nevertheless, if $Q_0 - \Delta Q$ is indefinite then, depending on R , the lower bounds in (9.6) and (9.7) may still be positive and thus be meaningful lower bounds.

10. Analytical Examples

In this section we consider simple analytical examples which illustrate the principal results of the paper. These examples also provide insight into the individual characteristics of different bounds as a prelude to numerical examples considered in the following section.

To begin we consider the simplest possible example. Set $n = 1$, $A < 0$, $R > 0$, $V > 0$, $A_1 = 1$, and $\mathcal{U} = \{\Delta A : |\Delta A| \leq \delta_1\}$. For $\delta_1 < -A$, $Q_{\Delta A} = V/2(|A| - \Delta A)$ and $J_S(\mathcal{U}) = J_D(\mathcal{U}) = RV/2(|A| - \delta_1)$, where this worst-case performance is achieved for $\Delta A = \delta_1$. Solving MLE1 yields $Q = V/2(|A| - \delta_1)$ which is a nonconservative result for both robust stability and performance. The same result is obtained from MLE4 by setting $\alpha = \alpha_1 = \delta_1$. To apply MLE5, set $\delta_1 = \sqrt{DE}$. Choosing $\alpha = 2\delta_1(|A| - \delta_1)EV$ again yields the nonconservative result. Finally, the same result follows from Theorem 8.1.

For the second example we consider nondestabilizing uncertainty in the imaginary component of an uncertain eigenvalue, i.e., frequency uncertainty, in contrast to uncertainty in the real part considered in the previous example. Let $n = 2$, $A = \begin{bmatrix} -\nu & \omega \\ \omega & -\nu \end{bmatrix}$, $\nu > 0$, $\omega \geq 0$, $V = R = I_2$, and $\mathcal{U} = \{\Delta A : \Delta A = \sigma_1 A_1, |\sigma_1| \leq \delta_1\}$, where $A_1 = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$. Obviously, $A + \Delta A$ remains asymptotically stable for all values of σ_1 since ΔA affects only the imaginary part of the poles of ΔA . The question then is whether the robustness tests are able to guarantee this robustness. Note also that because of the choice of V , $Q_{\Delta A} = Q_0 = (2\nu)^{-1}I_2$ for all $\Delta A \in \mathcal{U}$.

For this example we note that MLE1 is satisfied by $Q = (2\nu)^{-1}I_2$ which is independent of δ_1 . Thus MLE1 possesses a nonnegative-definite solution for all $\delta_1 > 0$ which shows that MLE1 is nonconservative with respect to robust stability and performance. Since $A(\Delta A) = (\Delta A)A$, it can also be seen that the same result holds for (8.18). The situation is considerably different for MLE4 and MLE5. To analyze MLE4 note that A has an eigenvalue $-2\nu + \alpha + \delta_1$. (This can be shown by diagonalizing A and A_1 and thus A .) Since, by Proposition 7.1, A must be asymptotically stable, we require $\delta_1 < 2\nu$. This is, of course, an extremely conservative result, especially when the damping ν is small. For MLE5 we can factor $A_1 = D_1 E_1$. Thus, let $D_1 = I_2$ and $E_1 = A_1$ and define $D = \delta_1^2 I_2$ and $E = I_2$. Assuming that Q is a multiple of I_2 , it follows that Q is nonnegative definite only if $\delta_1 \leq \nu$, which is again an extremely conservative result. The reason for this conservatism becomes clear by noting that D and E as given above will also serve as bounds for perturbations of the form $\sigma_1 I_2$ for which the range of nondestabilizing σ_1 is $|\sigma_1| < \delta_1$. This will also be the case for all factorizations $D_1 E_1$ of A_1 since $D_1 D_1^T$ and $E_1^T E_1$ must be positive definite and thus will also serve as bounds for destabilizing perturbations such as $\sigma_1 I_2$.

Finally, we consider a nondestabilizing uncertainty affecting the interaction of a pair of real poles. Let $n = 2$, $A = -I_2$, $V = R = I_2$, and $\mathcal{U} = \{\Delta A : \Delta A = \sigma_1 A_1, |\sigma_1| \leq \delta_1\}$, where $A_1 = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$. Obviously, $A + \Delta A$ remains stable for all values of σ_1 since ΔA does not affect the nominal poles. Note that $Q_{\Delta A} = \begin{bmatrix} \sigma_1^{2/4+1/2} & \sigma_1^{1/4} \\ \sigma_1^{1/4} & 1/2 \end{bmatrix}$ and $J_S(\mathcal{U}) = \frac{1}{4}\delta_1^2 + 1$, where this worst-case performance is achieved for $\sigma_1 = \delta_1$. In this case MLE1 has the solution $Q = (2 - \delta_1)^{-1} I_2$ which is valid only for $\delta_1 < 2$, an extremely conservative robust stability result. Furthermore, the corresponding performance bound $\text{tr } QR = 2(2 - \delta_1)^{-1}$ is conservative with respect to the actual worst-case performance $\frac{1}{4}\delta_1^2 + \frac{1}{2}$. In contrast MLE4 has the solution $Q = \begin{bmatrix} (2 - \alpha\delta_1)^{-1} + \alpha^{-1}\delta_1(2 - \alpha\delta_1)^{-2} & 0 \\ 0 & (2 - \alpha\delta_1)^{-1} \end{bmatrix}$ which is nonnegative definite for all δ_1 so long as $\alpha < 2/\delta_1$. Hence MLE4 is nonconservative with respect to robust stability. For robust performance, $\text{tr } QR = 2(2 - \alpha\delta_1)^{-1} + \alpha^{-1}\delta_1(2 - \alpha\delta_1)^{-2}$ which is clearly an upper bound for $\frac{1}{4}\delta_1^2 + 1$. Choosing, for example, $\alpha = \delta_1^{-1}$ yields $\text{tr } QR = \delta_1^2 + 2$. The parameter α can also be chosen to minimize $\text{tr } QR$, although this is somewhat tedious to carry out analytically. Finally, MLE5 has the solution $Q = \begin{bmatrix} \frac{1}{2}(1 + \alpha^{-1}\delta_1) & 0 \\ 0 & [1 - (1 - \alpha\delta_1)^{1/2}]/\alpha\delta_1 \end{bmatrix}$ which exists so long as $\alpha \leq 1/\delta_1$. Hence MLE5 is also nonconservative with respect to robust stability. Choosing $\alpha = 1/\delta_1$ yields $\text{tr } QR = \frac{1}{2}\delta_1^2 + \frac{3}{2}$ which lies above the minimal bound $\frac{1}{4}\delta_1^2 + 1$. Again, α can be chosen to minimize $\text{tr } QR$.

11. Numerical Example

In this section we consider additional examples illustrating the results developed in earlier sections. In contrast to the analytical examples considered in Section 10, however, we consider more complex examples by numerically solving the modified Lyapunov equations. Here we focus on MLE4 and MLE5 which are the easiest to solve numerically. Specifically, we solved MLE4 by using the representation (7.2) (although this may not be practical when n is large), and we solved MLE5 by means of a standard Riccati package. To simplify matters we consider only uncertainties ΔA of the form $\sigma_1 A_1$. Presentation and evaluation of robust stability and performance results for multiparameter uncertainty can be fairly complex and thus is deferred to a future numerical study.

Since both robustness tests MLE4 and MLE5 depend upon an arbitrary positive constant α , it is desirable to determine the value of α which yields the tightest (i.e., lowest) performance bound for each robust stability range. To this end we performed a simple one-dimensional search to determine the best such α . Although analytical techniques may assist in determining optimal values of α more efficiently, the search technique proved to be adequate for the examples considered here.

As a first example we consider the control system given in [1] to demonstrate the lack of a guaranteed gain margin for LQG controllers. Hence consider

$$\dot{x}_0(t) = A_0 x_0(t) + B_0 u(t) + w_1(t), \quad (11.1)$$

$$y(t) = C_0 x_0(t) + w_2(t), \quad (11.2)$$

with controller

$$\dot{x}_c(t) = A_c x_c(t) + B_c y(t), \quad (11.3)$$

$$u(t) = C_c x_c(t), \quad (11.4)$$

and performance

$$J = \lim_{t \rightarrow \infty} \mathbb{E}[x_0^T(t) R_1 x_0(t) + u^T(t) R_2 u(t)]. \quad (11.5)$$

The data are

$$A_0 = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}, \quad B_0 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \quad C_0 = [1 \quad 0],$$

$$V_1 = R_1 = \rho \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}, \quad V_2 = R_2 = 1,$$

where V_1 and V_2 are the intensities of $w_1(t)$ and $w_2(t)$, respectively. Uncertainty ΔB_0 in B_0 is thus represented by $\sigma_1 B_1$, where $B_1 = [0 \ 1]^T$. Thus the closed-loop system corresponds to

$$A = \begin{bmatrix} A_0 & B_0 C_e \\ B_e C_0 & A_e \end{bmatrix}, \quad A_1 = \begin{bmatrix} 0 & B_1 C_e \\ 0 & 0 \end{bmatrix},$$

$$R = \begin{bmatrix} R_1 & 0 \\ 0 & 0 \end{bmatrix}, \quad V = \begin{bmatrix} V_1 & 0 \\ 0 & B_e V_2 B_e^T \end{bmatrix},$$

where the zero in the (2,2) block of R denotes the fact that we are considering the robust performance bound for the state regulation cost only. Choosing $\rho = 60$, it follows that the LQG gains are given by

$$A_e = \begin{bmatrix} -9 & 1 \\ -20 & -9 \end{bmatrix}, \quad B_e = \begin{bmatrix} 10 \\ 10 \end{bmatrix}, \quad C_e = [-10 \quad -10].$$

For this controller the actual stability region corresponds to $\sigma_1 \in (-.07, .01)$ so that the largest symmetric region about $\sigma_1 = 0$ is $|\sigma_1| \leq .01$. The worst-case performance over each stability region $|\sigma_1| \leq \delta_1$ is denoted by the solid line in Figure 1, while the performance bounds obtained from MLE4 and MLE5 are shown for several values of δ_1 . For MLE5 we set $D_1 = [0 \ 1 \ 0 \ 0]^T$ and $E_1 = [0 \ 0 \ C_e]$. Note that MLE5 yields considerably tighter estimates of worst-case performance, particularly as δ_1 approaches .01. For MLE4, optimal values of α were in the range .0012 to .0058, while for MLE5 (with $\Omega'_{10}(\cdot)$, see (5.26)), α was in the range .0143 to .0020.

As a second example we consider a pair of nominally uncoupled oscillators with uncertain coupling. This example was considered in [45] using the majorant Lyapunov technique. Let

$$A = \begin{bmatrix} -\nu & \omega_1 & 0 & 0 \\ -\omega_1 & -\nu & 0 & 0 \\ 0 & 0 & -\nu & \omega_2 \\ 0 & 0 & -\omega_2 & -\nu \end{bmatrix}, \quad A_1 = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix},$$

$$\nu = .2, \quad \omega_1 = .2, \quad \omega_2 = 1.8, \quad R = V = I_4,$$

and, for MLE5, define $D = A_1$ and $E_1 = I_4$. We consider bounds on $J_S(\mathcal{U})$ only.

Figure 2 illustrates the exact worst case performance along with performance bounds obtained from MLE4 and MLE5. For MLE4 optimal values of α ranged from .036 to .141 while for MLE5 optimal α was between .361 and .096. Although MLE4 was slightly less conservative than MLE5, both bounds were able to guarantee robust stability only for $\delta_1 = .15$ while the largest stability region is actually $\delta_1 = .54$. It is interesting to contrast this result with [45] where the majorant Lyapunov technique yielded a robust stability range of $\delta_1 = .4$ for a richer class of off-diagonal blocks having maximum singular value less than δ_1 .

Conclusion. It seems clear that no single quadratic Lyapunov bound is superior to the others. Although the conservatism of each bound is problem dependent, it is desirable to better understand the nature of the conservatism in order to utilize the bounds in an effective manner. Finally, the example illustrated in Figure 2 may indicate fundamental limitations of the quadratic Lyapunov function approach to robustness as compared to the majcrant technique of [45]. These remain questions for future research.

Acknowledgment. We wish to thank A. W. Daubendiek for producing the numerical results in Section 11.

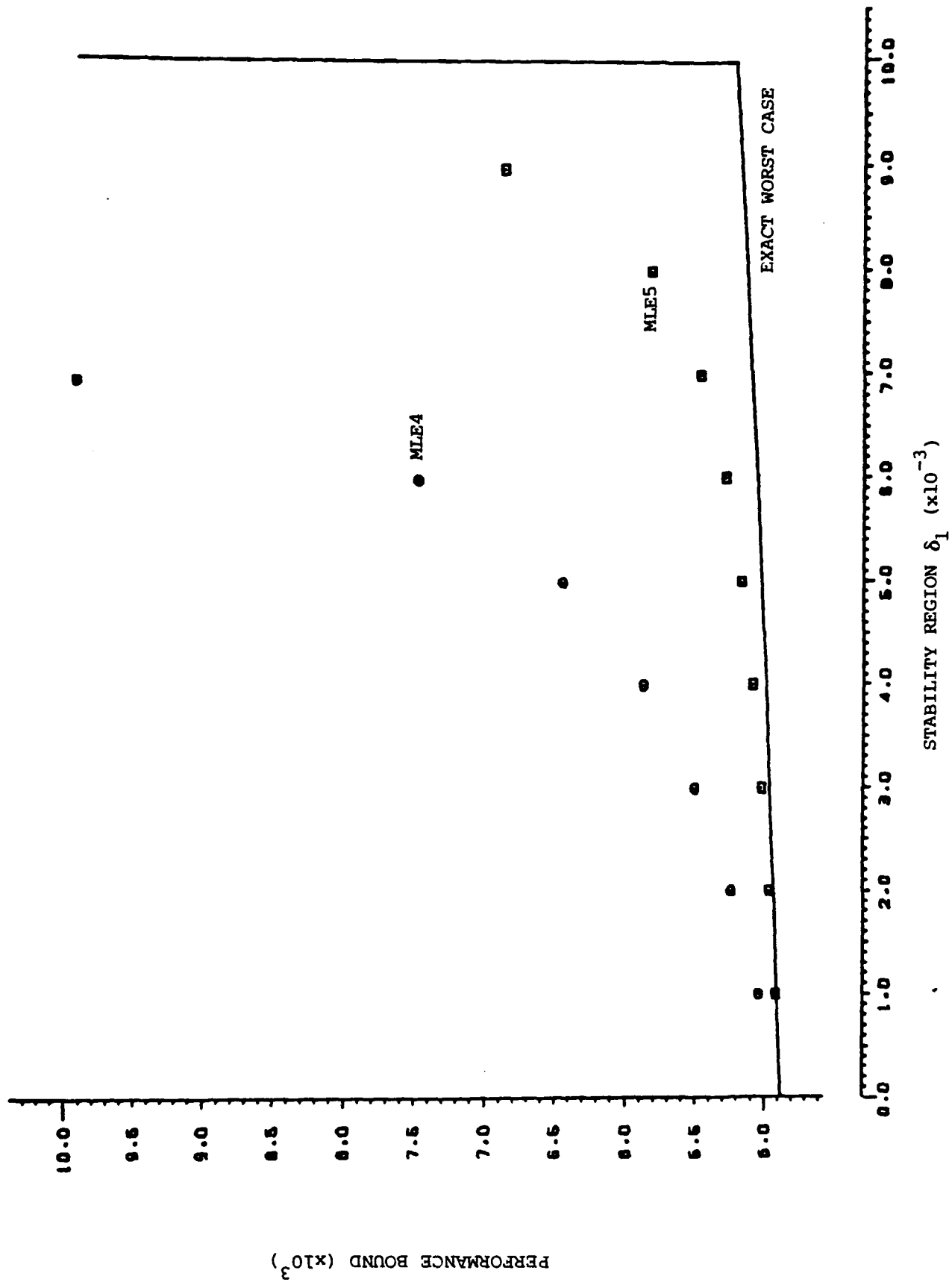


Figure 1

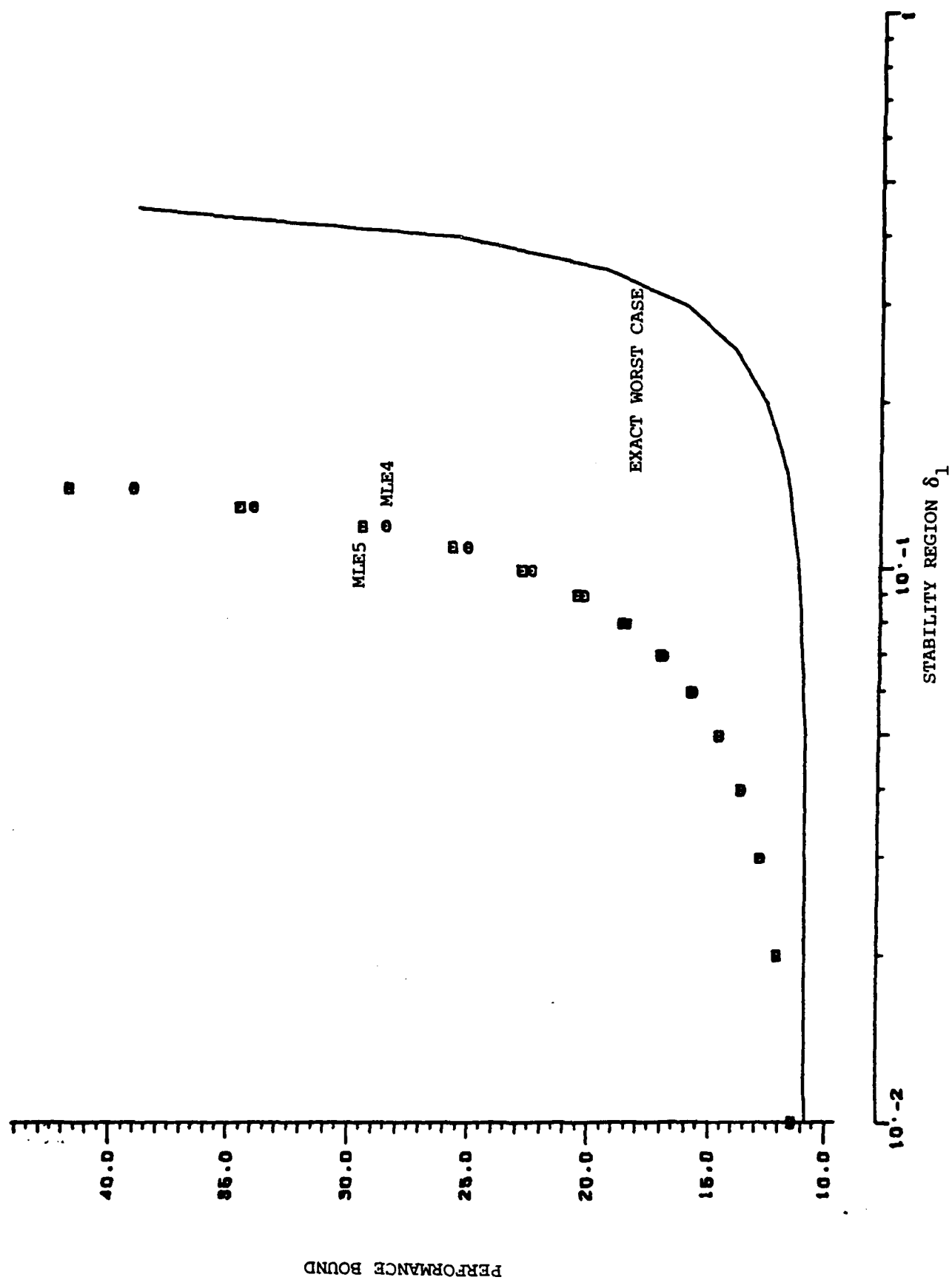


Figure 2

References

1. J. C. Doyle, "Guaranteed Margins for LQG Regulators," *IEEE Trans. Autom. Contr.*, Vol. AC-23, pp. 756-757, 1978.
2. E. Soroka and U. Shaked, "On the Robustness of LQ Regulators," *IEEE Trans. Autom. Contr.*, Vol. AC-29, pp. 664-665, 1984.
3. C. Van Loan, "How Near is a Stable Matrix to an Unstable Matrix?," *Contemporary Mathematics*, Vol. 47, pp. 465-478, American Mathematical Society, 1985.
4. D. Hinrichsen and A. J. Pritchard, "Stability Radii of Linear Systems," *Sys. Contr. Lett.*, Vol. 7, pp. 1-10, 1986.
5. J. M. Martin and G. A. Hewer, "Smallest Destabilizing Perturbations for Linear Systems," *Int. J. Contr.*, Vol. 45, pp. 1495-1504, 1987.
6. B. A. Francis, *A Course in H_∞ Control Theory*, Springer-Verlag, New York, 1987.
7. D. D. Siljak, "Parameter Space Methods for Robust Control Design: A Survey," preprint.
8. G. S. Michael and C. W. Merriam, "Stability of Parametrically Disturbed Linear Optimal Control Systems," *J. Math. Anal. Appl.*, Vol. 28, pp. 294-302, 1969.
9. S. S. L. Chang and T. K. C. Peng, "Adaptive Guaranteed Cost Control of Systems with Uncertain Parameters," *IEEE Trans. Autom. Contr.*, Vol. AC-17, pp. 474-483, 1972.
10. R. V. Patel, M. Toda and B. Sridhar, "Robustness of Linear Quadratic State Feedback Designs in the Presence of System Uncertainty," *IEEE Trans. Autom. Contr.*, Vol. AC-22, pp. 945-949, 1977.
11. G. Leitmann, "Guaranteed Asymptotic Stability for a Class of Uncertain Linear Dynamical Systems," *J. Optim. Thy. Appl.*, Vol. 27, pp. 99-106, 1979.
12. A. Vinkler and L. J. Wood, "Multistep Guaranteed Cost Control of Linear Systems with Uncertain Parameters," *J. Guid. Contr.*, Vol. 2, pp. 449-456, 1979.
13. M. Eslami and D. L. Russell, "On Stability with Large Parameter Variations Stemming from the Direct Method of Lyapunov," *IEEE Trans. Autom. Contr.*, Vol. AC-25, pp. 1231-1234, 1980.
14. R. V. Patel and M. Toda, "Quantitative Measures of Robustness for Multivariable Systems," in *Proc. Joint. Autom. Contr. Conf.*, paper TP8-A, San Francisco, CA, June 1980.
15. M. Corless and G. Leitmann, "Continuous State Feedback Guaranteeing Uniform Ultimate Boundedness for Uncertain Dynamical Systems," *IEEE Trans. Autom. Contr.*, Vol. AC-26, pp. 1139-1144, 1981.
16. J. S. Thorp and B. R. Barmish, "On Guaranteed Stability of Uncertain Linear Systems via Linear Control," *J. Optim. Thy. Appl.*, Vol. 35, pp. 559-579, 1981.
17. E. Noldus, "Design of Robust State Feedback Laws," *Int. J. Contr.*, Vol. 35, pp. 935-944, 1982.
18. B. R. Barmish, M. Corless, and G. Leitmann, "A New Class of Stabilizing Controllers for Uncertain Dynamical Systems," *SIAM J. Contr. Optim.*, Vol. 21, pp. 246-255, 1983.
19. B. R. Barmish, I. R. Petersen, and A. Feuer, "Linear Ultimate Boundedness Control of Uncertain Dynamic Systems," *Automatica*, Vol. 19, pp. 523-532, 1983.
20. B. R. Barmish, "Necessary and Sufficient Conditions for Quadratic Stabilizability of an Uncertain Linear System," *J. Optim. Thy. Appl.*, Vol. 46, pp. 399-408, 1985.
21. R. K. Yedavalli, S. S. Banda, and D. B. Ridgely, "Time-Domain Stability Robustness Measures for Linear Regulators," *J. Guid. Contr. Dyn.*, Vol. 8, pp. 520-524, 1985.
22. R. K. Yedavalli, "Improved Measures of Stability Robustness for Linear State Space Models," *IEEE Trans. Autom. Contr.*, Vol. AC-30, pp. 577-579, 1985.
23. D. S. Bernstein and D. C. Hyland, "The Optimal Projection/Maximum Entropy Approach to Designing Low-Order, Robust Controllers for Flexible Structures," *Proc. 24th IEEE Conf.*

- Dec. Contr.*, pp. 745-752, Fort Lauderdale, FL, December 1985.
24. A. R. Galimidi and B. R. Barmish, "The Constrained Lyapunov Problem and its Application to Robust Output Feedback Stabilization," *IEEE Trans. Autom. Contr.*, Vol. AC-31, pp. 410-419, 1986.
 25. I. R. Petersen and C. V. Hollot, "A Riccati Equation Approach to the Stabilization of Uncertain Systems," *Automatica*, Vol. 22, pp. 397-411, 1986.
 26. D. S. Bernstein and S. W. Greeley, "Robust Output-Feedback Stabilization: Deterministic and Stochastic Perspectives," *Proc. Amer. Contr. Conf.*, pp. 1818-1826, Seattle, WA, June 1986.
 27. D. S. Bernstein and S. W. Greeley, "Robust Controller Synthesis Using the Maximum Entropy Design Equations," *IEEE Trans. Autom. Contr.*, Vol. AC-31, pp. 362-364, 1986.
 28. I. R. Petersen, "A Stabilization Algorithm for a Class of Uncertain Linear Systems," *Sys. Contr. Lett.*, Vol. 8, pp. 351-357, 1987.
 29. D. Hinrichsen and A. J. Pritchard, "Stability Radii and Structured Perturbations and the Algebraic Riccati Equation," *Sys. Contr. Lett.*, Vol. 8, pp. 105-113, 1987.
 30. D. S. Bernstein and W. M. Haddad, "The Optimal Projection Equations with Petersen-Hollot Bounds: Robust Controller Synthesis with Guaranteed Structured Stability Radius," *Proc. IEEE Conf. Dec. Contr.*, pp. 1308-1318, Los Angeles, CA, December 1987.
 31. K. Zhou and P. P. Khargonekar, "Stability Robustness Bounds for Linear State-Space Models with Structured Uncertainty," *IEEE Trans. Autom. Contr.*, Vol. AC-32, pp. 621-623, 1987.
 32. S. P. Bhattacharyya, *Robust Stabilization Against Structured Perturbations*, Springer-Verlag, 1987.
 33. D. S. Bernstein, "Robust Static and Dynamic Output-Feedback Stabilization: Deterministic and Stochastic Perspectives," *IEEE Trans. Autom. Contr.*, Vol. AC-32, pp. 1076-1084, 1987.
 34. O. I. Kosmidou and P. Bertrand, "Robust-Controller Design for Systems with Large Parameter Variations," *Int. J. Contr.*, Vol. 45, pp. 927-938, 1987.
 35. D. S. Bernstein and W. M. Haddad, "Robust Stability and Performance for Fixed-Order Dynamic Compensation via the Optimal Projection Equations with Guaranteed Cost Bounds," *Amer. Contr. Conf.*, Atlanta, GA, June 1988.
 36. D. S. Bernstein and W. M. Haddad, "The Optimal Projection Equations with Petersen-Hollot Bounds: Robust Stability and Performance via Fixed-Order Dynamic Compensation for Systems with Structured Real-Valued Parameter Uncertainty," *IEEE Trans. Autom. Contr.*, Vol. AC-33, pp. 578-582, 1988.
 37. D. S. Bernstein and W. M. Haddad, "Robust Stability and Performance Analysis for Linear Dynamic Systems," submitted.
 38. D. A. Wilson, "Convolution and Hankel Operator Norms for Linear Systems," *IEEE Trans. Autom. Contr.*, to appear.
 39. W. M. Wonham, *Linear Multivariable Control: A Geometric Approach*, Springer-Verlag, New York, 1979.
 40. R. A. Horn and C. R. Johnson, *Matrix Analysis*, Cambridge Univ. Press, 1985.
 41. T. Kato, *Perturbation Theory for Linear Operators*, Springer-Verlag, 1980.
 42. J. W. Brewer, "Kronecker Products and Matrix Calculus in System Theory," *IEEE Trans. Circ. Sys.*, Vol. CAS-25, pp. 772-781, 1978.
 43. L. Arnold, *Stochastic Differential Equations: Theory and Applications*, Wiley, New York, 1974.
 44. E. G. Collins, Jr., and D. C. Hyland, "Improved Robust Performance Bounds in Covariance Majorant Analysis," preprint.
 45. D. C. Hyland and D. S. Bernstein, "The Majorant Lyapunov Equation: A Nonnegative Matrix Equation for Guaranteed Robust Stability and Performance of Large Scale Systems," *IEEE Trans. Autom. Contr.*, Vol. AC-32, pp. 1005-1013, 1987.

The Majorant Lyapunov Equation: A Nonnegative Matrix Equation for Robust Stability and Performance of Large Scale Systems

DAVID C. HYLAND AND DENNIS S. BERNSTEIN, MEMBER, IEEE

Abstract—A new robust stability and performance analysis technique is developed. The approach involves replacing the state covariance by its block-norm matrix, i.e., the nonnegative matrix whose elements are the norms of subblocks of the covariance matrix partitioned according to subsystem dynamics. A bound (i.e., majorant) for the block-norm matrix is given by the majorant Lyapunov equation, a Lyapunov-type nonnegative matrix equation. Existence, uniqueness, and computational tractability of solutions to the majorant Lyapunov equation are shown to be completely characterized in terms of M matrices. Two examples are considered. For a damped simple harmonic oscillator with uncertain but constant natural frequency, the majorant Lyapunov equation predicts unconditional stability. And, for a pair of nominally uncoupled oscillators with uncertain coupling, the majorant Lyapunov equation shows that the range of nondestabilizing couplings is proportional to the frequency separation between the oscillators, a result not predictable from quadratic or vector Lyapunov functions.

I. INTRODUCTION

THE importance of robustness in control-system analysis and design cannot be overemphasized. The past ten years' literature reflects considerable frequency-domain development [1]–[11], while recent publications indicate increasing time-domain activity [12]–[19]. Wide variations in underlying assumptions, mathematical settings, and problem data render it difficult, if not impossible, to clearly delineate the relative effectiveness of different methods. Our own philosophical outlook has thus been guided by two general criteria:

- 1) effectiveness for simple examples;
- 2) efficiency when applied to large scale problems.

The first criterion involves applying robustness techniques to simple, perhaps trivially obvious, examples to serve as "acid tests." A given method's effectiveness on a collection of such examples can possibly reveal inherent shortcomings. As an illustration of this criterion, consider a damped harmonic oscillator with constant but uncertain natural frequency. Using the notation of [6], stability is guaranteed so long as

$$\sigma_{\max}[R(j\omega)(I + G(j\omega)K(j\omega))^{-1}G(j\omega)L^{-1}(j\omega)] < 1, \quad \omega \geq 0 \quad (1.1)$$

where, for $\nu > 0$,

$$G(s) = (s^2 + 2\nu s + \omega_n^2)^{-1}$$

Manuscript received August 8, 1986; revised May 6, 1987. Paper recommended by Associate Editor, M. G. Safonov. This work was supported in part by the Air Force Office of Scientific Research under Contracts F49620-86-C-0002 and F49620-86-C-0038.

The authors are with the Government Aerospace Systems Division, Harris Corporation, Melbourne, FL 32902.
IEEE Log Number 8716541.

and uncertainty in the nominal natural frequency ω_n is modeled by

$$\begin{aligned} \Delta(s) &= L^{-1}(s)\theta(s)R(s) = \delta\omega_n^2, \\ L(s) &= 1/\alpha, \quad \theta(s) = \delta/\alpha, \quad R(s) = \omega_n^2, \quad K(s) = 0, \\ \delta &\in [-\min(1, \alpha), \alpha], \quad \alpha > 0. \end{aligned}$$

Note that

$$\sigma_{\max}[\theta(j\omega)] \leq 1, \quad \omega \geq 0$$

as required in [6]. The perturbation $\Delta(s)$ (modeled as a feedback gain) effectively replaces ω_n^2 in $G(s)$ by $(1 + \delta)\omega_n^2$. Hence, for a given $\alpha > 0$ this uncertainty model permits perturbed natural frequencies in the range $[0, (1 + \alpha)^{-1/2}\omega_n]$. Evaluating (1.1) yields the upper bound

$$\alpha < [(\omega_n^2 - \omega^2)^2 + 4\nu^2\omega^2]^{1/2}/\omega_n^2, \quad \omega \geq 0 \quad (1.2)$$

or, equivalently,

$$\alpha < 2\zeta(1 - \zeta^2)^{1/2} \quad (1.3)$$

where $\zeta \triangleq \nu/\omega_n$. The conservatism of (1.3) is obviously most pronounced when the damping ratio ζ is small. In all cases, however, the conservatism is infinite.

The second criterion is obviously subjective and depends upon a variety of factors such as problem structure, designer experience, and computational resources. This criterion is, in our opinion, most important since the need for robustness techniques becomes increasingly critical as system complexity grows. Indeed, the ultimate test of a given approach is to scale it up to larger and larger problems to reveal inherent limitations. Obviously, such tests are not only difficult, but may entail a significant commitment of human and financial resources. Nevertheless, crude predictions are sometimes available, and a case in point is the "curse of dimensionality" encountered in the approach of [9]. Another example involves computational difficulties in obtaining bounds for the μ -function with more than three blocks [10].

The contribution of the present paper is a new robustness analysis method developed specifically for large scale systems. The basic idea, motivated by the work of Siljak [30] on connective stability, is as follows. The system is assumed to be in the form of a collection of subsystems with uncertain local dynamics and uncertain interactions.¹ Parameter uncertainties are modeled as either structured or unstructured constant variations contained in prescribed sets. The state covariance, partitioned conformably with the subsystem dynamics, is replaced by its *block-norm matrix*, i.e., the nonnegative matrix each of whose elements is the norm of the corresponding subblock of the original matrix. This nonnegative matrix satisfies a novel inequality designated the

¹ Uncertainties in a single subsystem can also be regarded as interaction uncertainties. To see this, write $\dot{x} = (A + G)x$ twice so that the uncertainty G is represented by $\begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$.

covariance block-norm inequality. The existence of a solution to the *majorant Lyapunov equation*, i.e., the covariance block-norm inequality interpreted as an *equation*, yields an element-by-element bound (i.e., majorant) for the covariance block-norm, hence, assuring robust stability and performance. The relevance of this technique to large scale systems stems from the fact that replacing each subblock of the covariance by its norm can significantly reduce the dimension of the problem. Indeed, the dimension of the majorant Lyapunov equation is equal to the number of subsystems which may be significantly less than the dimension of the original system.

To illustrate the above ideas in more detail, consider the covariance equation

$$0 = (\bar{A} + G)Q + Q(\bar{A} + G)^T + V \quad (1.4)$$

where \bar{A} denotes the nominal dynamics, G denotes uncertainty in \bar{A} , V is the disturbance intensity, and Q is the state covariance. Assuming that \bar{A} is block diagonal with r diagonal blocks leads to the covariance block-norm inequality (see Proposition 4.2)

$$\alpha * Q \leq \leq GQ + QG^T + V. \quad (1.5)$$

In (1.5), α , Q , G , and V are $r \times r$ nonnegative matrices, i.e., each element is a nonnegative number. The matrices Q and V are formed by taking the Frobenius norm of each subblock of Q and V , while each component of G is a given constant which bounds the spectral norm (largest singular value) of the corresponding subblock of the uncertain perturbation G . Hence, G is a majorant for G in the sense of [21]–[23]. Each element of the matrix α is bounded above by the smallest singular value of the Kronecker sum [24]–[26] of pairs of diagonal blocks of \bar{A} . The operation “*” is the Hadamard product [27], [28], and the ordering “ \leq ” denotes element-by-element comparison, i.e., the ordering induced by the cone of nonnegative matrices [29], [30].

The majorant Lyapunov equation is obtained by replacing the inequality (1.5) by the $r \times r$ nonnegative matrix equation

$$\alpha * \tilde{Q} = G\tilde{Q} + \tilde{Q}G^T + V. \quad (1.6)$$

A key result (Corollary 5.1) states that

$$Q \leq \tilde{Q} \quad (1.7)$$

for all stable $\bar{A} + G$. Consequently (see Theorem 5.1), the existence of a unique solution to (1.6) leads directly to a guarantee of robust stability over the range specified by G and to a performance bound involving \tilde{Q} . Moreover, solutions of (1.6) exist if and only if the $r^2 \times r^2$ matrix

$$\bar{A} \triangleq \text{diag}(\text{vec } \alpha) - G \otimes G \quad (1.8)$$

is an M matrix [29], [30].

Even when the number of subsystems is large, the majorant Lyapunov equation is generally computationally tractable. Specifically, although \bar{A} is an $r^2 \times r^2$ matrix, *no computations whatsoever need to be carried out with matrices of this dimension*. Rather, it suffices to solve only the majorant Lyapunov equation (1.6). In this regard we show that \tilde{Q} is given by

$$\tilde{Q} = \lim_{i \rightarrow \infty} \tilde{Q}_i \quad (1.9)$$

where if G has only off-diagonal nonzero blocks the sequence $\{\tilde{Q}_i\}$ is generated by

$$\alpha * \tilde{Q}_{i+1} = G\tilde{Q}_i + \tilde{Q}_iG^T + V, \quad \tilde{Q}_0 = 0 \quad (1.10)$$

and is monotonically increasing. Furthermore, the convergence of this sequence is equivalent to \bar{A} being an M matrix so that it is not even necessary to form \bar{A} . Note that (1.6) does not require the development of new solution techniques. Indeed, since (1.10) is a

straightforward iteration, (1.6) is even easier to solve than the original Lyapunov equation (1.4).

To illustrate these results we consider two examples. The first example is the damped oscillator already considered in this section. With little effort the majorant Lyapunov equation yields the (obvious) result that the oscillator is stable for *all* constant natural frequencies. The second example involves a pair of oscillators with known parameters but with uncertain coupling. The majorant Lyapunov equation yields bounds over which stability is guaranteed, and these bounds are compared to the actual stability region as a function of frequency separation. The main result shows that the robustness to uncertain coupling is proportional to the frequency separation. This weak subsystem interaction robustification mechanism is the principal contribution of the majorant theory. This example has immediate application to the problem of vibration control in flexible structures. For this class of problems the open-loop dynamics can be viewed as a collection of uncoupled oscillators which become coupled via feedback and structural uncertainties.

The majorant bound developed in the present paper is quite different from the widely used quadratic Lyapunov function (see, e.g., [12]–[20]). As can readily be shown using the methods of [12], [17]–[20], the quadratic Lyapunov function yields robust stability and performance by replacing (1.4) by

$$0 = \bar{A}\tilde{Q} + \tilde{Q}\bar{A}^T + \Omega(\tilde{Q}) + V \quad (1.11)$$

where $\Omega(\cdot)$ satisfies

$$G\tilde{Q} + \tilde{Q}G^T \leq \Omega(\tilde{Q}) \quad (1.12)$$

for all variations G . It can then be shown that

$$Q \leq \tilde{Q} \quad (1.13)$$

where now, in contrast to (1.7), the ordering in (1.13) is defined with respect to the cone of *nonnegative-definite* matrices. Indeed, the majorant bound may be more closely related to vector Lyapunov functions [30], [31] and the Lyapunov matrix function [32], [33]. It does not appear possible, however, to use these techniques to obtain the majorant results on robustness due to subsystem frequency separation.

The reader will observe that this paper exploits a wide variety of techniques including nonnegative matrices, block norms, matrix majorants, the Hadamard product, the Kronecker sum, and M matrices. Each of these techniques, except majorants, has, however, been previously applied to control problems in numerous instances. In the special case of scalar subblocks, the block-norm matrix has, moreover, been utilized by Yedavalli [13]–[15] and others for robustness analysis and design. In this case the block norm is known as the matrix *modulus*. The variety of algebraic structures employed in the present paper should not be surprising since the quest for increasingly refined robustness techniques can be expected to invoke correspondingly refined uncertainty bounds. Related techniques are employed in [11]. Furthermore, nonnegative matrix equations involving M matrices arise naturally in a variety of settings (see, e.g., [38], [39]).

The contents of the paper are as follows. Section II presents notation, definitions, and lemmas for use throughout the paper. In Section III robust stability and performance are defined for the homogeneous and nonhomogeneous systems. Detailed system structure and uncertainty characterization are given in Section IV and the covariance block-norm inequality is derived. Section V analyzes the majorant Lyapunov equation to obtain a majorant for the steady-state covariance. The main result, Theorem 5.1, guarantees robust stability and provides a performance bound. Finally, the examples appear in Section VI.

II. PRELIMINARIES

The following notation will be used throughout. All matrices are assumed to have real entries.

| | |
|---|--|
| \mathbb{R} | expected value |
| $\mathbb{R}, \mathbb{R}^{p \times q}, \mathbb{R}^p$ | real numbers, $p \times q$ real matrices, $\mathbb{R}^{p \times 1}$ |
| $I_p, 0_{p \times q}, 0_p$ | $p \times p$ identity matrix, $p \times q$ zero matrix, $0_{p \times p}$ |
| \odot, \otimes | Kronecker sum, Kronecker product [24]–[27] |
| * | Hadamard product [27], [28] |
| $\text{col}_i(Z)$ | i th column of matrix Z |
| $\text{vec}(Z)$ | $\begin{bmatrix} \text{col}_1(Z) \\ \vdots \\ \text{col}_q(Z) \end{bmatrix} \in \mathbb{R}^n, Z \in \mathbb{R}^{p \times q}$ |
| $Z_{(i,j)}$ | (i, j) element of matrix Z |
| Z^T | transpose of vector or matrix Z |
| Z^{-T} | $(Z^T)^{-1}$ or $(Z^{-1})^T$ |
| $\text{tr } Z$ | trace of matrix Z |
| $\text{diag}(Z_1, \dots, Z_p)$ | diagonal matrix with listed diagonal elements |
| $\text{block-diag}(Z_1, \dots, Z_p)$ | block-diagonal matrix with listed diagonal blocks |
| $\rho(Z)$ | spectral radius of Z |
| asymptotically stable matrix | matrix with eigenvalues in open left-half plane |
| nonnegative-definite matrix | symmetric matrix with nonnegative eigenvalues ($Z \geq 0$) |
| positive-definite matrix | symmetric matrix with positive eigenvalues ($Z > 0$) |
| $Z_1 \geq Z_2$ | $Z_1 - Z_2 \geq 0$, Z_1, Z_2 symmetric |
| $Z_1 > Z_2$ | $Z_1 - Z_2 > 0$, Z_1, Z_2 symmetric |
| nonnegative matrix | matrix with nonnegative elements ($Z \geq 0$) [29], [30] |
| positive matrix | matrix with positive elements ($Z > 0$) |
| $Z_1 \geq Z_2$ | $Z_1 - Z_2 \geq 0$ |
| $Z_1 > Z_2$ | $Z_1 - Z_2 > 0$ |
| Z^{HI} | Hadamard inverse, $(Z^{HI})_{(i,j)} \triangleq [Z_{(i,j)}]^{-1}, Z > 0$ |
| block-norm matrix | nonnegative matrix each of whose elements is the norm of a corresponding subblock of a given partitioned matrix |
| majorant | nonnegative matrix each of whose elements bounds the corresponding element of a block-norm matrix |
| $\ Z\ _2$ | Euclidean norm of vector Z |
| $\sigma_i(Z)$ | singular value of matrix Z |
| $\sigma_{\min}(Z), \sigma_{\max}(Z)$ | smallest and largest singular values of matrix Z |
| $\lambda_{\max}(Z)$ | largest eigenvalue of symmetric matrix Z |
| $\ Z\ _s$ | $\sigma_{\max}(Z)$ (spectral norm induced by $\ \cdot\ _2$) |
| $\ Z\ _F$ | $(\text{tr } ZZ^T)^{1/2} = \left[\sum_{i,j=1}^{p,q} Z_{(i,j)}^2 \right]^{1/2} = \left[\sum_{i=1}^p \sigma_i^2 \right]^{1/2}$ (Frobenius norm [34]). |

In subsequent sections we shall exploit the fact that the norms $\|\cdot\|_2$, $\|\cdot\|_s$, and $\|\cdot\|_F$ coincide for vectors. Hence, if $Z \in \mathbb{R}^p$, then by interpreting $\mathbb{R}^p = \mathbb{R}^{p \times 1}$ it follows that

$$\|Z\|_2 = \|Z\|_s = \|Z\|_F. \quad (2.1)$$

Furthermore, if $Z \in \mathbb{R}^{p \times q}$, then

$$\|Z\|_s \leq \|Z\|_F = \|\text{vec } Z\|_F = \|\text{vec } Z\|_2 = \|\text{vec } Z\|_s. \quad (2.2)$$

Lemma 2.1: If $Z \in \mathbb{R}^{p \times q}$ and $\hat{Z} \in \mathbb{R}^{q \times r}$ then

$$\sigma_{\min}(Z) \| \hat{Z} \|_F \leq \| Z \hat{Z} \|_F \leq \| Z \|_s \| \hat{Z} \|_F. \quad (2.3, 2.4)$$

If, furthermore, $p = q = r$, $Z \geq 0$, and \hat{Z} is symmetric, then

$$\text{tr } Z \hat{Z} \leq (\text{tr } Z) \lambda_{\max}(\hat{Z}) \leq (\text{tr } Z) \| \hat{Z} \|_s. \quad (2.5)$$

Proof: Inequality (2.4) can be found in [35, p. 263]. To prove (2.3), note that when Z is singular the result is immediate. Otherwise, if $p = q$ replace Z and \hat{Z} in (2.4) by Z^{-1} and $Z \hat{Z}$, respectively. The result now follows from $[\sigma_{\max}(Z)]^{-1} = \sigma_{\min}(Z^{-1})$. If $p \neq q$, then related arguments apply. Finally, (2.5) is given in [36]. \square

Recall [30] that a matrix $S \in \mathbb{R}^{r \times r}$ is an N matrix if $S_{(i,j)} \leq 0$, $i, j = 1, \dots, r$, $i \neq j$. If, in addition, all principal minors of S are positive, then S is an M matrix.

Lemma 2.2: Suppose $S \in \mathbb{R}^{r \times r}$ is an N matrix. Then the following are equivalent:

- S is an M matrix;
- $\det S \neq 0$ and $S^{-1} \geq 0$;
- for each $y \in \mathbb{R}^r$, $y \geq 0$, there exists a unique $x \in \mathbb{R}^r$, $x \geq 0$, such that $Sx = y$;
- there exists $x \in \mathbb{R}^r$, $x \geq 0$, such that $Sx > 0$;
- $I_r + S > 0$ and each diagonal matrix $D > I_r + S$ satisfies $\rho[D^{-1}(I_r + S - S)] < 1$.

Proof: The equivalence of statements i), ii), iv), and v) follows from [30, p. 396]. The implication ii) \Rightarrow iii) is immediate, and iii) \Rightarrow iv) follows by setting $y = [1 \ 1 \ \dots \ 1]^T$. \square

Lemma 2.3: Suppose $S \in \mathbb{R}^{r \times r}$ is an M matrix and let $\hat{S} \in \mathbb{R}^{r \times r}$ be an N matrix such that $\hat{S} \geq S$. Then \hat{S} is an M matrix.

Proof: See [30, p. 400]. \square

III. ROBUST STABILITY AND PERFORMANCE BOUNDS

Consider the n th-order homogeneous system²

$$\dot{x}(t) = (A(\theta) + G)x(t), \quad t \in [0, \infty), \quad (3.1)$$

$$G \in \mathfrak{G} \subset \mathbb{R}^{n \times n}, \quad (3.2)$$

$$\theta \in \Theta \subset \mathbb{R}^m, \quad (3.3)$$

where $A: \Theta \rightarrow \mathbb{R}^{n \times n}$ is continuous, $\bar{A} \triangleq A(\bar{\theta})$ denotes the known nominal dynamics for $\bar{\theta} \in \Theta$, θ denotes the unstructured parametric uncertainty in \bar{A} , G denotes the structured parametric uncertainty in \bar{A} , and $0 \in \mathfrak{G}$ is the nominal value of G . We first consider the stability of (3.1) over \mathfrak{G} and Θ .

Definition 3.1: If $A(\theta) + G$ is asymptotically stable for all $G \in \mathfrak{G}$ and $\theta \in \Theta$, then the homogeneous system (3.1) is *robustly stable* over \mathfrak{G} and Θ .

Now consider the n th-order nonhomogeneous system

$$\dot{x}(t) = (A(\theta) + G)x(t) + w(t), \quad t \in [0, \infty) \quad (3.4)$$

where $G \in \mathfrak{G}$, $\theta \in \Theta$, and $w(\cdot)$ is white noise with intensity $V \geq 0$. For given $G \in \mathfrak{G}$ and $\theta \in \Theta$, the steady-state average quadratic performance is defined by

$$J(G, \theta) \triangleq \limsup_{t \rightarrow \infty} \mathbb{E}[x^T(t) R x(t)] \quad (3.5)$$

where $R = R^T \geq 0$. The system (3.4) may, for example, denote a control system in closed-loop configuration. There is no need in our development, however, to make such distinctions.

In practice, steady-state performance is only of interest when the system is robustly stable. The following result is immediate.

Proposition 3.1: Suppose the system (3.1) is robustly stable

² Upon first reading the uncertainty represented by (3.3) can be ignored since the principal contribution concerns the treatment of (3.2).

over \mathfrak{G} and Θ . Then for each $G \in \mathfrak{G}$ and $\theta \in \Theta$,

$$J(G, \theta) = \text{tr } Q R \quad (3.6)$$

where $n \times n$ nonnegative-definite Q is the unique solution to

$$0 = (A(\theta) + G)Q + Q(A(\theta) + G)^T + V. \quad (3.7)$$

We shall only be concerned with the case in which \mathfrak{G} and Θ are compact. Since Q is a continuous function of G and θ , we can define the worst-case average steady-state quadratic performance

$$J_{\max} \triangleq \max_{G \in \mathfrak{G}, \theta \in \Theta} J(G, \theta). \quad (3.8)$$

Since it is difficult to determine J_{\max} explicitly, we shall seek upper bounds.

Definition 3.2: If $J_{\max} \leq \hat{\alpha}$, then $\hat{\alpha}$ is a *performance bound* for the nonhomogeneous system (3.4) over \mathfrak{G} and Θ .

IV. SYSTEM STRUCTURE, UNCERTAINTY CHARACTERIZATION, AND THE COVARIANCE BLOCK-NORM INEQUALITY

A discussed in Section I, (3.1) and (3.4) are assumed to be in the form of a large scale system with uncoupled local dynamics and uncertain interactions. Hence, with the subsystem partitioning

$$n = \sum_{i=1}^r n_i \quad (4.1)$$

the local system dynamics $A(\theta)$ can be decomposed into subsystem dynamics according to

$$A(\theta) = \text{block-diag} \{A_i(\theta)\}_{i=1, \dots, r} \quad (4.2)$$

where $A_i(\theta) \in \mathbb{R}^{n_i \times n_i}$, $\theta \in \Theta$. For convenience, denote

$$\bar{A} \triangleq \text{block-diag} \{\bar{A}_i\}_{i=1, \dots, r}$$

Accordingly, R is assumed to be of the form

$$R = \text{block-diag} \{R_i\}_{i=1, \dots, r} \quad (4.3)$$

where $R_i \in \mathbb{R}^{n_i \times n_i}$, $R_i \geq 0$, $i = 1, \dots, r$. The intensity V and steady-state covariance Q satisfying (3.7) are assumed to be conformably partitioned, i.e.,

$$V = \{V_{ij}\}_{i,j=1}^r, \quad V_{ij} \in \mathbb{R}^{n_i \times n_j}, \quad (4.4)$$

$$Q = \{Q_{ij}\}_{i,j=1}^r, \quad Q_{ij} \in \mathbb{R}^{n_i \times n_j}. \quad (4.5)$$

For notational simplicity define

$$V_i \triangleq V_{ii}, \quad Q_i \triangleq Q_{ii}, \quad i = 1, \dots, r. \quad (4.6)$$

Taking the Frobenius norm of each subblock of V and Q leads to the $r \times r$ symmetric nonnegative matrices \mathcal{V} and \mathcal{Q} defined by

$$\mathcal{V} \triangleq \{\|V_{ij}\|_F\}_{i,j=1}^r, \quad \mathcal{Q} \triangleq \{\|Q_{ij}\|_F\}_{i,j=1}^r. \quad (4.7)$$

Note that

$$\|\mathcal{Q}\|_F = \|Q\|_F, \quad \|\mathcal{V}\|_F = \|V\|_F. \quad (4.8)$$

A few observations concerning the nominal system, i.e., with $G = 0$ and $\theta = \bar{\theta}$, are worth noting. If \bar{A} is stable then so is \bar{A}_i , $i = 1, \dots, r$, and there exist unique, nonnegative-definite \bar{Q}_i , $\bar{P}_i \in \mathbb{R}^{n_i \times n_i}$, $i = 1, \dots, r$, satisfying

$$0 = \bar{A}_i \bar{Q}_i + \bar{Q}_i \bar{A}_i^T + V_i, \quad (4.9)$$

$$0 = \bar{A}_i^T \bar{P}_i + \bar{P}_i \bar{A}_i + R_i. \quad (4.10)$$

Proposition 4.1: Suppose \bar{A} is asymptotically stable. Then the nominal performance J_{nom} is given by

$$J_{\text{nom}} \triangleq J(0, \bar{\theta}) = \sum_{i=1}^r \text{tr } \bar{Q}_i R_i = \sum_{i=1}^r \text{tr } \bar{P}_i V_i. \quad (4.11)$$

Proof: First note that with $G = 0$ and $\theta = \bar{\theta}$ the diagonal blocks of Q satisfying (3.7) coincide with $\bar{Q}_1, \dots, \bar{Q}_r$. Thus

$$\begin{aligned} J(0, \bar{\theta}) &= \sum_{i=1}^r \text{tr } \bar{Q}_i R_i \\ &= \sum_{i=1}^r (\text{vec } \bar{Q}_i)^T \text{vec } R_i \\ &= \sum_{i=1}^r [(\bar{A}_i \oplus \bar{A}_i)^{-1} \text{vec } V_i]^T \text{vec } R_i \\ &= \sum_{i=1}^r (\text{vec } V_i)^T (\bar{A}_i^T \oplus \bar{A}_i^T)^{-1} \text{vec } R_i \\ &= \sum_{i=1}^r (\text{vec } V_i)^T \text{vec } \bar{P}_i \\ &= \sum_{i=1}^r \text{tr } \bar{P}_i V_i. \quad \square \end{aligned}$$

The matrices $G \in \mathfrak{G}$ are also conformably partitioned so that

$$G = \{G_{ij}\}_{i,j=1}^r, \quad G_{ij} \in \mathbb{R}^{n_i \times n_j} \quad (4.12)$$

and \mathfrak{G} is characterized by

$$\mathfrak{G} \triangleq \{G \in \mathbb{R}^{n \times n} : \sigma_{\max}(G_{ij}) \leq \gamma_{ij}, \quad i, j = 1, \dots, r\} \quad (4.13)$$

where $\gamma_{ij} \geq 0$, $i, j = 1, \dots, r$, are given constants. For convenience, define the $r \times r$ nonnegative matrix

$$\mathcal{G} \triangleq \{\gamma_{ij}\}_{i,j=1}^r. \quad (4.14)$$

The bound \mathcal{G} is a *matrix majorant* for $G \in \mathfrak{G}$ in the sense of [21]–[23].

Remark 4.1: \mathfrak{G} is compact and convex.

Finally, let symmetric, positive $\mathcal{Q} \in \mathbb{R}^{r \times r}$ satisfy

$$\mathcal{Q}_{(i,j)} \leq \min_{\theta \in \Theta} \{\sigma_{\min}(A_j(\theta) \oplus A_i(\theta))\}, \quad i, j = 1, \dots, r. \quad (4.15)$$

Proposition 4.2: Let $G \in \mathfrak{G}$ and $\theta \in \Theta$ be such that $A(\theta) + G$ is asymptotically stable and let $n \times n$ $Q \geq 0$ satisfy (3.7). Then \mathcal{Q} defined by (4.7) satisfies

$$\mathcal{Q} * \mathcal{Q} \leq \mathcal{G} \mathcal{Q} + \mathcal{Q} \mathcal{G}^T + \mathcal{V} \quad (4.16)$$

or, equivalently,

$$\bar{\mathcal{A}} \text{vec } \mathcal{Q} \leq \text{vec } \mathcal{V}, \quad (4.17)$$

where

$$\bar{\mathcal{A}} \triangleq [\text{diag}(\text{vec } \mathcal{Q})] - \mathcal{G} \oplus \mathcal{G}. \quad (4.18)$$

Proof: Expanding (3.7) yields

$$\begin{aligned} -[A_i(\theta) Q_{ij} + Q_{ij} A_j^T(\theta)] &= \sum_{k=1}^r [G_{ik} Q_{kj} + Q_{ik} G_{jk}^T] + V_{ij}, \\ i, j &= 1, \dots, r. \end{aligned} \quad (4.19)$$

Bounding the right-hand side of (4.19) from above using (2.4) yields for all $G \in \mathfrak{G}$

$$\left\| \sum_{k=1}^r [G_{ik} Q_{kj} + Q_{ik} G_{jk}^T] + V_{ij} \right\|_F \leq \sum_{k=1}^r [\mathfrak{G}_{(i,k)} \mathfrak{Q}_{(k,j)} + \mathfrak{Q}_{(i,k)} \mathfrak{G}_{(j,k)}] + \mathfrak{V}_{(i,j)}$$

while bounding the left-hand side of (4.19) from below using (2.3) implies for all $\theta \in \Theta$

$$\begin{aligned} \left\| -[A_i(\theta) Q_{ij} + Q_{ij} A_j^T(\theta)] \right\|_F &= \left\| \text{vec}(A_i(\theta) Q_{ij} + Q_{ij} A_j^T(\theta)) \right\|_F \\ &= \left\| (A_i(\theta) \otimes A_i(\theta)) \text{vec } Q_{ij} \right\|_F \\ &\geq \sigma_{\min}(A_j(\theta) \otimes A_i(\theta)) \left\| \text{vec } Q_{ij} \right\|_F \\ &= \sigma_{\min}(A_j(\theta) \otimes A_i(\theta)) \mathfrak{Q}_{(i,j)} \\ &\geq \mathfrak{Q}_{(i,j)} \mathfrak{Q}_{(i,j)}. \end{aligned}$$

Combining the above inequalities yields (4.16). \square

Remark 4.2: Since $\mathfrak{G} \geq 0$, the $r^2 \times r^2$ matrix \mathfrak{A} is an N matrix [30].

V. THE MAJORANT LYAPUNOV EQUATION

In this section we interpret (4.16) as an equality rather than an inequality and consider the Lyapunov-type nonnegative matrix equation

$$\mathfrak{A} * \tilde{\mathfrak{Q}} = \mathfrak{G} \tilde{\mathfrak{Q}} + \tilde{\mathfrak{Q}} \mathfrak{G}^T + \mathfrak{V} \quad (5.1)$$

or, equivalently,

$$\mathfrak{A} \text{vec } \tilde{\mathfrak{Q}} = \text{vec } \mathfrak{V}. \quad (5.2)$$

Note that since \mathfrak{A} and \mathfrak{V} are symmetric a unique solution of (5.1) is necessarily symmetric.

Proposition 5.1: The following are equivalent:

- \mathfrak{A} is an M matrix;
- $\det \mathfrak{A} \neq 0$ and $\mathfrak{A}^{-1} \geq 0$;
- for each $r \times r$ symmetric $\mathfrak{V} \geq 0$ there exists a unique $r \times r$ $\tilde{\mathfrak{Q}} \geq 0$ satisfying (5.1);
- there exist $r \times r$ symmetric $\mathfrak{V} \gg 0$ and $r \times r$ symmetric $\tilde{\mathfrak{Q}} \geq 0$ satisfying (5.1);
- $\text{diag}(\text{vec } \mathfrak{A}) - (I, * \mathfrak{G}) \otimes (I, * \mathfrak{G}) \gg 0$ and each diagonal matrix $\mathfrak{D} \geq \text{diag}(\text{vec } \mathfrak{A}) - (I, * \mathfrak{G}) \otimes (I, * \mathfrak{G})$ satisfies

$$\rho(\mathfrak{D}^{-1}[\mathfrak{G} \otimes \mathfrak{G} - (I, * \mathfrak{G}) \otimes (I, * \mathfrak{G})]) < 1; \quad (5.3)$$

- for each $r \times r$ symmetric $\tilde{\mathfrak{Q}}_0 \geq 0$ and $r \times r$ symmetric $\mathfrak{V} \geq 0$, the sequence $\{\tilde{\mathfrak{Q}}_i\}_{i=1}^\infty$ generated by

$$\begin{aligned} \mathfrak{Q} * \tilde{\mathfrak{Q}}_{i+1} - (I, * \mathfrak{G}) \tilde{\mathfrak{Q}}_{i+1} - \tilde{\mathfrak{Q}}_{i+1} (I, * \mathfrak{G}) \\ = (\mathfrak{G} - I, * \mathfrak{G}) \tilde{\mathfrak{Q}}_i + \tilde{\mathfrak{Q}}_i (\mathfrak{G} - I, * \mathfrak{G})^T + \mathfrak{V}, \quad i=0, 1, \dots, \end{aligned} \quad (5.4)$$

converges;

- for each $r \times r$ symmetric $\tilde{\mathfrak{Q}}_0 \geq 0$ there exists $r \times r$ symmetric $\mathfrak{V} \gg 0$ such that the sequence $\{\tilde{\mathfrak{Q}}_i\}_{i=1}^\infty$ generated by (5.4) converges.

Proof: Statements i)–v) are equivalent to i)–v) of Lemma 2.2. Clearly, vi) implies iii), and vii) implies iv). To show v) implies vi) and vii) note that $I, * (\mathfrak{G} \otimes \mathfrak{G}) = (I, * \mathfrak{G}) \otimes (I, * \mathfrak{G})$ and

$$\text{vec}(\mathfrak{A} * \tilde{\mathfrak{Q}}_{i+1}) = [\text{diag}(\text{vec } \mathfrak{A})] \text{vec } \tilde{\mathfrak{Q}}_{i+1}.$$

Thus, (5.4) is equivalent to

$$\begin{aligned} \text{vec } \tilde{\mathfrak{Q}}_{i+1} &= [\text{diag}(\text{vec } \mathfrak{A}) - (I, * \mathfrak{G}) \otimes (I, * \mathfrak{G})]^{-1} \\ &\quad \cdot [\mathfrak{G} \otimes \mathfrak{G} - I, * (\mathfrak{G} \otimes \mathfrak{G})] \text{vec } \tilde{\mathfrak{Q}}_i + [\text{diag}(\text{vec } \mathfrak{A})]^{-1} \text{vec } \mathfrak{V}. \end{aligned}$$

Thus, vi) and vii) follow from v) with $\mathfrak{D} = \text{diag}(\text{vec } \mathfrak{A}) - (I, * \mathfrak{G}) \otimes (I, * \mathfrak{G})$. \square

Since statements i)–vii) depend only upon \mathfrak{A} and \mathfrak{G} we have the following definition inspired by v)–vii).

Definition 5.1: $(\mathfrak{A}, \mathfrak{G})$ is *stable* if \mathfrak{A} is an M matrix.

Remark 5.1: When $I, * \mathfrak{G} = 0$, i.e., when the local dynamics have no structured uncertainty, (5.4) simplifies to

$$\mathfrak{A} * \tilde{\mathfrak{Q}}_{i+1} = \mathfrak{G} \tilde{\mathfrak{Q}}_i + \tilde{\mathfrak{Q}}_i \mathfrak{G}^T + \mathfrak{V}, \quad i=0, 1, \dots, \quad (5.5)$$

or, equivalently,

$$\tilde{\mathfrak{Q}}_{i+1} = \mathfrak{A}^{H^i} * (\mathfrak{G} \tilde{\mathfrak{Q}}_i + \tilde{\mathfrak{Q}}_i \mathfrak{G}^T + \mathfrak{V}), \quad i=0, 1, \dots. \quad (5.5a)$$

The following result shows that for zero initial condition, the iterative sequence is monotonic.

Proposition 5.2: Suppose $\text{diag}(\text{vec } \mathfrak{A}) - I, * (\mathfrak{G} \otimes \mathfrak{G}) \gg 0$. Then the sequence $\{\tilde{\mathfrak{Q}}_i\}_{i=1}^\infty$ generated by (5.4) with $\tilde{\mathfrak{Q}}_0 = 0$ and $\mathfrak{V} \geq 0$ is monotonically increasing.

Proof: To simplify notation we consider the case mentioned in Remark 5.1. Hence, assume $\mathfrak{A} \gg 0$. Clearly, if $\tilde{\mathfrak{Q}}_0 = 0$, then (5.5a) implies that $\tilde{\mathfrak{Q}}_1 = \mathfrak{A}^{H^1} * \mathfrak{V} \geq 0$. Hence, $\tilde{\mathfrak{Q}}_1 \geq \tilde{\mathfrak{Q}}_0$. Defining $\Delta \tilde{\mathfrak{Q}}_{i+1} \triangleq \tilde{\mathfrak{Q}}_{i+1} - \tilde{\mathfrak{Q}}_i$, (5.5a) yields

$$\Delta \tilde{\mathfrak{Q}}_{i+1} = \mathfrak{A}^{H^i} * (\mathfrak{G} \Delta \tilde{\mathfrak{Q}}_i + \Delta \tilde{\mathfrak{Q}}_i \mathfrak{G}^T).$$

Since $\Delta \tilde{\mathfrak{Q}}_1 \geq 0$, the result follows from induction. \square

Remark 5.2: Proposition 5.2 is a particularly useful result in applications and can be utilized as follows. Setting $\tilde{\mathfrak{Q}}_0 = 0$, the sequence $\{\tilde{\mathfrak{Q}}_i\}$ can be evaluated by a simple numerical procedure. As will be shown in Theorem 5.1 below, each $\tilde{\mathfrak{Q}}_i$ corresponds to a robust performance measure $\hat{\alpha}_i$. For practical purposes the increasing sequence $\{\hat{\alpha}_i\}$ can be generated until either convergence is attained (in which case $\hat{\alpha} = \lim_{i \rightarrow \infty} \hat{\alpha}_i$ is a robust performance bound) or a maximum permissible performance level is exceeded. In the latter case the question of convergence is irrelevant since the closed-loop system is known to either be unstable for some $G \in \mathfrak{G}$ (i.e., $\hat{\alpha} = \infty$) or exceed acceptable performance specifications, thereby necessitating system redesign.

We now prove a comparison result for solutions of (5.1).

Lemma 5.1: Assume $(\mathfrak{A}, \mathfrak{G})$ is stable, let $\hat{\mathfrak{A}}, \hat{\mathfrak{G}}$ be $r \times r$ nonnegative matrices where $\hat{\mathfrak{A}}$ is symmetric, and assume that

$$\mathfrak{A} \leq \hat{\mathfrak{A}}, \quad \mathfrak{G} \leq \hat{\mathfrak{G}}. \quad (5.6)$$

Then $(\hat{\mathfrak{A}}, \hat{\mathfrak{G}})$ is stable. Furthermore, let $r \times r$ symmetric \mathfrak{V} satisfy

$$\mathfrak{V} \leq \hat{\mathfrak{V}}, \quad (5.7)$$

let $\tilde{\mathfrak{Q}}$ be the unique, nonnegative solution to (5.1), and let $\hat{\tilde{\mathfrak{Q}}}$ be the unique solution to

$$\hat{\mathfrak{A}} * \hat{\tilde{\mathfrak{Q}}} = \hat{\mathfrak{G}} \hat{\tilde{\mathfrak{Q}}} + \hat{\tilde{\mathfrak{Q}}} \hat{\mathfrak{G}}^T + \hat{\mathfrak{V}}. \quad (5.8)$$

Then if $\tilde{\mathfrak{Q}} \geq 0$, it follows that

$$\hat{\tilde{\mathfrak{Q}}} \leq \tilde{\mathfrak{Q}}. \quad (5.9)$$

Proof: Since

$$\hat{\mathfrak{A}} \triangleq \text{diag}(\text{vec } \hat{\mathfrak{A}}) - \hat{\mathfrak{G}} \otimes \hat{\mathfrak{G}}$$

is an N matrix, $\hat{\mathfrak{A}}$ is an M matrix, and

$$\hat{\mathfrak{A}} - \mathfrak{A} = \text{diag}(\text{vec } (\hat{\mathfrak{A}} - \mathfrak{A})) + (\mathfrak{G} - \hat{\mathfrak{G}}) \otimes (\mathfrak{G} - \hat{\mathfrak{G}}) \geq 0$$

it follows from Lemma 2.3 that \hat{A} is an M matrix, and thus (\hat{Q}, \hat{G}) is stable. Next note that (5.1) and (5.8) imply

$$\text{vec}(\hat{Q} - \tilde{Q}) = \hat{A}^{-1}(\hat{A} - A) \text{vec} \hat{Q} + \hat{A}^{-1} \text{vec}(\Psi - \tilde{\Psi}).$$

Since $\hat{A} - A \geq 0$, $\hat{A}^{-1} \geq 0$ (see Lemma 2.2), $\Psi - \tilde{\Psi} \geq 0$, and $\hat{Q} \geq 0$, it follows that (5.9) is satisfied. \square

Corollary 5.1: Suppose (\hat{Q}, \hat{G}) is stable and let \tilde{Q} be the unique, nonnegative solution to (5.1). Furthermore, let $G \in \mathfrak{G}$ and $\theta \in \Theta$ be such that $A(\theta) + G$ is asymptotically stable and define Q by (4.7) for $n \times n$ $Q \geq 0$ satisfying (3.7). Then

$$Q \leq \tilde{Q}. \quad (5.10)$$

Proof: By Proposition 4.2, Q satisfies the covariance block-norm inequality (4.16). In the notation of Lemma 5.1 define

$$\hat{Q} = \hat{Q}, \hat{G} = \hat{G}, \hat{\Psi} = \hat{Q} * \hat{Q} - (\hat{G}\hat{Q} + \hat{Q}\hat{G}^T) \quad (5.11)$$

so that (5.6) is satisfied and (4.16) implies (5.7). Note that with the notation (5.11), equation (5.8) has the unique solution $\hat{Q} = \tilde{Q} \geq 0$. Hence (5.9) implies (5.10). \square

Theorem 5.1: Assume A is asymptotically stable, Θ is continuously arcwise connected, and (\hat{Q}, \hat{G}) is stable. Then the homogeneous system (3.1) is robustly stable over \mathfrak{G} and Θ , and the nonhomogeneous system (3.4) has the performance bound

$$\hat{\alpha} = \max_{\theta \in \Theta} \left\{ \sum_{i=1}^r [\text{tr}(\hat{Q}_i(\theta)R_i) + 2(\text{tr} \hat{P}_i(\theta))(\hat{G}\hat{Q}_{(i,i)})] \right\} \quad (5.12)$$

where $n_i \times n_i$ nonnegative-definite $\hat{Q}_i(\theta)$ and $\hat{P}_i(\theta)$ satisfy

$$0 = A_i(\theta)\hat{Q}_i(\theta) + \hat{Q}_i(\theta)A_i^T(\theta) + V_i, \quad (5.13)$$

$$0 = A_i^T(\theta)\hat{P}_i(\theta) + \hat{P}_i(\theta)A_i(\theta) + R_i \quad (5.14)$$

and $r \times r$ \hat{Q} is the unique, nonnegative solution to (5.1).

Proof: First note that since robust stability is independent of the disturbances, we can set $V = I_n$ for convenience in proving the first result. Hence, suppose (3.1) is not robustly stable. Since \mathfrak{G} is convex (see Remark 4.1), A is asymptotically stable, and Θ is continuously arcwise connected, there exist $G_0 \in \mathfrak{G}$ and $\hat{\theta}: [0, 1] \rightarrow \Theta$ such that $\hat{A}(\mu) \triangleq A(\hat{\theta}(\mu)) + \mu G_0$ is asymptotically stable for all $\mu \in [0, 1]$, and $\hat{A}(1)$ is not asymptotically stable. Define

$$Q(\mu, t) \triangleq \int_0^t e^{\hat{A}(\mu)s} e^{\hat{A}^T(\mu)s} ds, \quad t \geq 0, \mu \in [0, 1]$$

which is monotonically increasing in the nonnegative-definite cone with respect to t . Clearly, the limit

$$Q(\mu) \triangleq \lim_{t \rightarrow \infty} Q(\mu, t), \quad \mu \in [0, 1]$$

exists and satisfies

$$0 = \hat{A}(\mu)Q(\mu) + Q(\mu)\hat{A}^T(\mu) + I_n, \quad \mu \in [0, 1].$$

Now define $r \times r$ nonnegative symmetric $Q(\mu)$ by

$$Q(\mu) = \{\|Q_{ij}(\mu)\|_F\}_{i,j=1}^r$$

where $Q_{ij}(\mu) \in \mathbb{R}^{n_i \times n_j}$ and $Q(\mu)$ is partitioned as in (4.5). By Corollary 5.1 with $\theta = \hat{\theta}(\mu)$, $G = \mu G_0$, $Q \triangleq Q(\mu)$, $\mu \in [0, 1]$, and $V = I_n$, it follows from (5.10) that

$$Q(\mu) \leq \tilde{Q}, \quad \mu \in [0, 1]. \quad (5.15)$$

Hence, by (4.8), (5.15) implies

$$\|Q(\mu)\|_F = \|Q(\mu)\|_F \leq \|\tilde{Q}\|_F, \quad \mu \in [0, 1]. \quad (5.16)$$

On the other hand, for $\mu \in [0, 1]$ it follows that

$$\begin{aligned} Q(\mu) &= Q(\mu) - Q(\mu, t) + Q(\mu, t) - Q(1, t) + Q(1, t) \\ &\geq Q(\mu, t) - Q(1, t) + Q(1, t) \end{aligned}$$

which implies, for arbitrary $x \in \mathbb{R}^n$,

$$x^T Q(\mu)x \geq x^T [Q(\mu, t) - Q(1, t)]x + x^T Q(1, t)x.$$

Thus, by continuity of $Q(\mu, t)$ in μ ,

$$\lim_{\mu \rightarrow 1} x^T Q(\mu)x \geq x^T Q(1, t)x, \quad x \in \mathbb{R}^n. \quad (5.17)$$

Now, since $\hat{A}(1)$ is not asymptotically stable and $(\hat{A}(1), I_n)$ is stabilizable, it follows from [37, Proposition 3.2, p. 67] that for some $\bar{x} \in \mathbb{R}^n$,

$$\lim_{t \rightarrow \infty} \bar{x}^T Q(1, t)\bar{x} = \infty.$$

Thus, by (5.17)

$$\lim_{\mu \rightarrow 1} \bar{x}^T Q(\mu)\bar{x} = \infty$$

and thus

$$\lim_{\mu \rightarrow 1} \|Q(\mu)\|_F = \infty. \quad (5.18)$$

However, (5.18) contradicts (5.16). Hence, (3.1) is robustly stable over \mathfrak{G} and Θ .

To derive (5.12) note that since R is block diagonal,

$$J(G, \theta) = \sum_{i=1}^r \text{tr} Q_i R_i = \sum_{i=1}^r (\text{vec} Q_i)^T \text{vec} R_i$$

where Q satisfies (3.7). Furthermore, (4.19) implies

$$\text{vec} Q_i = -[A_i(\theta) \oplus A_i(\theta)]^{-1} \cdot \left[\text{vec} V_i + \sum_{k=1}^r \text{vec} (G_{ik} Q_{ki} + Q_{ik} G_{ik}^T) \right].$$

Hence, using Lemma 2.1,

$$\begin{aligned} J(G, \theta) &= \sum_{i=1}^r \left[\text{tr}(\hat{Q}_i(\theta)R_i) \right. \\ &\quad \left. + \sum_{k=1}^r (\text{vec} [G_{ik} Q_{ki} + Q_{ik} G_{ik}^T])^T \text{vec} \hat{P}_i(\theta) \right] \\ &= \sum_{i=1}^r \left[\text{tr}(\hat{Q}_i(\theta)R_i) + \sum_{k=1}^r \text{tr} \hat{P}_i(\theta)(G_{ik} Q_{ki} + Q_{ik} G_{ik}^T) \right] \\ &\leq \sum_{i=1}^r \left[\text{tr}(\hat{Q}_i(\theta)R_i) \right. \\ &\quad \left. + \sum_{k=1}^r (\text{tr} \hat{P}_i(\theta)) \sigma_{\max}(G_{ik} Q_{ki} + Q_{ik} G_{ik}^T) \right] \\ &\leq \sum_{i=1}^r \left[\text{tr}(\hat{Q}_i(\theta)R_i) + 2(\text{tr} \hat{P}_i(\theta)) \sum_{k=1}^r \sigma_{\max}(G_{ik}) \sigma_{\max}(Q_{ki}) \right] \\ &\leq \sum_{i=1}^r \left[\text{tr}(\hat{Q}_i(\theta)R_i) + 2(\text{tr} \hat{P}_i(\theta)) \sum_{k=1}^r \sigma_{\max}(G_{ik}) \|Q_{ki}\|_F \right] \end{aligned}$$

$$\leq \sum_{i=1}^r \left[\text{tr} (\bar{Q}_i(\theta) R_i) + 2 (\text{tr} \bar{P}_i(\theta)) \sum_{k=1}^r G_{(i,k)} \bar{Q}_{(k,i)} \right]$$

$$= \sum_{i=1}^r [\text{tr} (\bar{Q}_i(\theta) R_i) + 2 (\text{tr} \bar{P}_i(\theta)) (G \bar{Q})_{(i,i)}]$$

which yields (5.12). \square

VI. EXAMPLES

We first confirm that the damped harmonic oscillator is asymptotically stable for all constant frequency perturbations. Hence, let

$$r=1, n=n_1=2$$

and

$$\bar{A} = \bar{A}_1 = \begin{bmatrix} -\nu & \omega \\ -\omega & -\nu \end{bmatrix}$$

where $\nu > 0$ and $\omega \in \mathbb{R}$. To represent frequency uncertainty let $\Theta = \{\theta\}$, $\Theta = \mathbb{R}$, $\bar{\theta} = 0$, and

$$A(\theta) = \bar{A} + \theta \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}.$$

Note that $A(\theta)$ is stable for all $\theta \in \mathbb{R}$ with poles $-\nu \pm j(\omega + \theta)$. Note that $A(\theta)$ can be diagonalized by means of the unitary transformation

$$\phi = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ j & -j \end{bmatrix}, \quad \phi^{-1} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & -j \\ 1 & j \end{bmatrix}$$

so that

$$\bar{A}(\theta) \triangleq \phi^{-1} A(\theta) \phi = \begin{bmatrix} -\nu + j(\omega + \theta) & 0 \\ 0 & -\nu - j(\omega + \theta) \end{bmatrix}.$$

Hence, using

$$A(\theta) \otimes A(\theta) = (\phi^{-1} \otimes \phi^{-1})(\bar{A}(\theta) \otimes \bar{A}(\theta))(\phi \otimes \phi)$$

it follows that

$$\sigma_{\min}(A(\theta) \otimes A(\theta)) = 2\nu, \quad \theta \in \mathbb{R}.$$

Defining [see (4.15)]

$$\bar{\alpha} = \bar{\alpha}_{(1,1)} = 2\nu$$

and $\bar{g} = 0$, the scalar majorant Lyapunov equation (5.1) has the solution

$$\bar{Q} = \bar{V}/2\nu$$

where $\bar{V} = \|V\|_F$. Choosing $V = I_2$ and noting that $\bar{A} = \bar{\alpha} = 2\nu > 0$ is an M matrix, Theorem 5.1 guarantees robust stability for all frequency variations $\theta \in \mathbb{R}$.

The next example has been chosen to demonstrate the robustness of a pair of nominally uncoupled oscillators with respect to uncertain coupling. Hence, let

$$n=4, r=2, n_1=n_2=2$$

and

$$\bar{A}_i = \begin{bmatrix} -\nu & \omega_i \\ -\omega_i & -\nu \end{bmatrix}, \quad i=1, 2$$

where $\nu, \omega_1, \omega_2 \geq 0$. Furthermore, let $\Theta = \{\theta\}$ and

$$G = \begin{bmatrix} 0 & \gamma_{12} \\ \gamma_{21} & 0 \end{bmatrix}$$

which denotes the fact that the local subsystem (oscillator) dynamics are assumed to be known. Since

$$\sigma_{\min}(\bar{A}_j \otimes \bar{A}_i) = [4\nu^2 + (\omega_j - \omega_i)^2]^{1/2}$$

define

$$\bar{\alpha} = \begin{bmatrix} 2\nu & [4\nu^2 + (\omega_1 - \omega_2)^2]^{1/2} \\ [4\nu^2 + (\omega_1 - \omega_2)^2]^{1/2} & 2\nu \end{bmatrix}.$$

Letting $V = I_4$ yields $\bar{V} = 2I_2$. Solving (5.1) yields

$$\bar{Q}_{(1,1)} = (2\nu^2\delta - \gamma_{12}\gamma_{21} + \gamma_{12}^2)/2\sqrt{2}(\nu^2\delta - \gamma_{12}\gamma_{21}),$$

$$\bar{Q}_{(1,2)} = (\gamma_{12} + \gamma_{21})/2\sqrt{2}(\nu^2\delta - \gamma_{12}\gamma_{21}),$$

$$\bar{Q}_{(2,2)} = (2\nu^2\delta - \gamma_{12}\gamma_{21} + \gamma_{21}^2)/2\sqrt{2}(\nu^2\delta - \gamma_{12}\gamma_{21})$$

where

$$\delta \triangleq [1 + \delta^2]^{1/2}, \quad \delta \triangleq (\omega_1 - \omega_2)/2\nu.$$

Clearly, \bar{Q} is nonnegative if and only if

$$\gamma_{12}\gamma_{21} < \nu^2\delta. \quad (6.1)$$

The bound (6.1) characterizes the magnitude of coupling uncertainty for which stability is guaranteed. Note that the parameter δ is a measure of the frequency separation of the oscillators relative to the damping. When $\delta \gg 1$, (6.1) becomes asymptotically

$$\gamma_{12}\gamma_{21} < \frac{\nu}{2} |\omega_1 - \omega_2| \quad (6.2)$$

which confirms the intuitive expectation that robust stability is proportional to damping and subsystem frequency separation. This result does not appear to be predictable from quadratic or vector Lyapunov functions.

To evaluate the conservatism inherent in the bound (6.1) we solve for the actual stability region. To render the calculation tractable we assume that G_{12} and G_{21} have the structured form

$$G_{ij} = \begin{bmatrix} \alpha_{ij} & \beta_{ij} \\ -\beta_{ij} & \alpha_{ij} \end{bmatrix}. \quad (6.3)$$

By constraining (6.3) the set of coupling variations is reduced, which may or may not lead to a larger stability region. Thus, our estimate of conservatism may itself be conservative, i.e., the actual conservatism may indeed be less than the following analysis indicates. However, without (6.3) the development becomes intractable. This calculation will thus be called *semiexact*.

By considering the characteristic equation for $\bar{A} + G$, lengthy manipulation shows that $\bar{A} + G$ is stable if and only if

$$\gamma_{12}\gamma_{21} < 2\nu^2(-\epsilon + [1 + \delta^2(1 - \epsilon^2)]^{1/2})/(1 - \epsilon^2) \quad (6.4)$$

where $\epsilon \in (0, 1]$ is the smallest positive real root of

$$\epsilon = (1 + \epsilon^2)[1 + \delta^2(1 - \epsilon^2)]^{1/2}/[2 + \delta^2(1 - \epsilon^2)]. \quad (6.5)$$

The majorant bound (6.1) and semiexact bound (6.4) are illustrated in unified form in Fig. 1. For $\delta \gg 1$ note that $\epsilon = O(\delta^{-1})$ and thus (6.4) becomes asymptotically

$$\gamma_{12}\gamma_{21} < \nu |\omega_1 - \omega_2|. \quad (6.6)$$

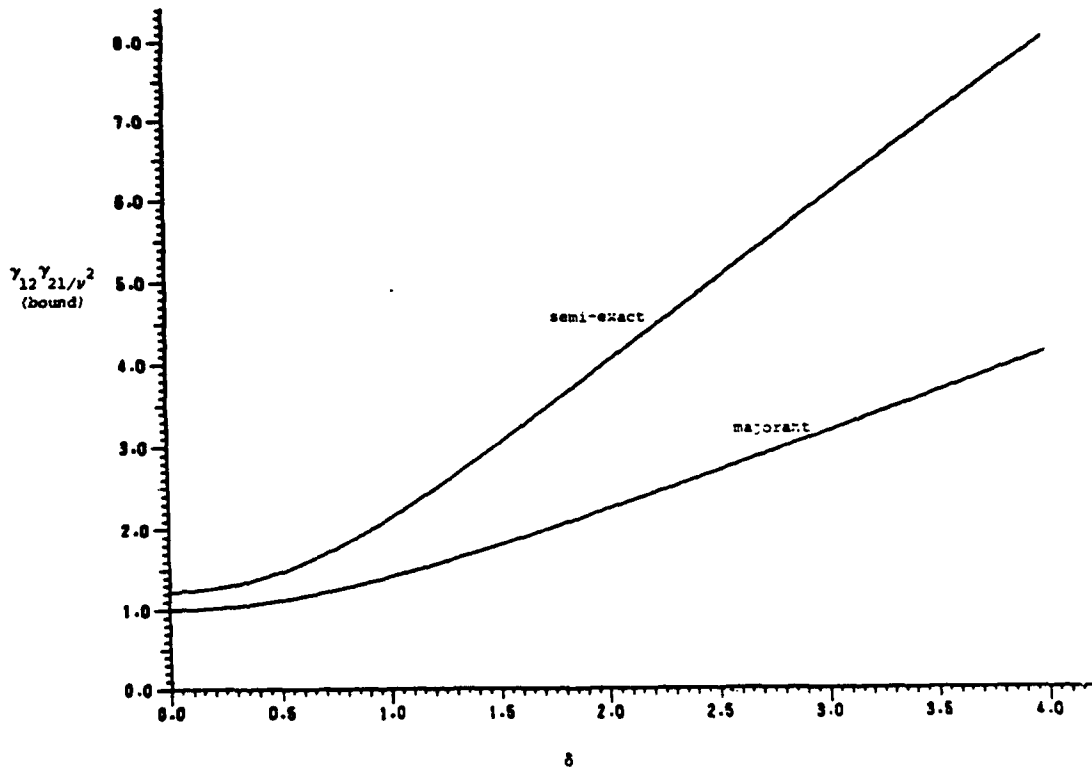


Fig. 1.

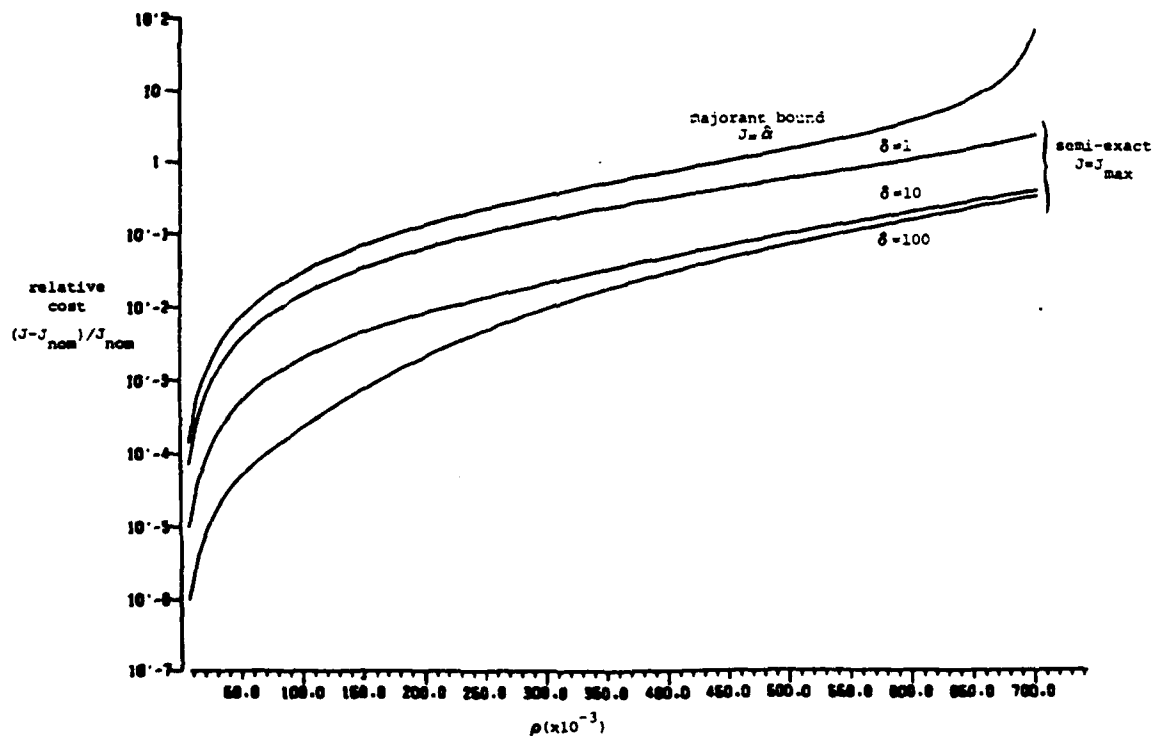


Fig. 2.

Hence, for large δ the majorant bound (6.2) is, at worst, conservative by a factor of 2 compared to the semiexact bound.

To determine the performance bound (5.12) set $R = I_4$. Hence, it can be shown that

$$J_{nom} = 2/\nu$$

and the system has the performance bound

$$\hat{\alpha} = J_{nom} + \sqrt{2}(\rho_{12} + \rho_{21})^2 / \nu(1 - 2\rho_{12}\rho_{21}) \quad (6.7)$$

where

$$\rho_{12} = \gamma_{12} / \sqrt{2\nu\delta^{1/2}}, \quad \rho_{21} = \gamma_{21} / \sqrt{2\nu\delta^{1/2}}.$$

On the other hand, the semiexact calculation yields

$$J_{\max} = \max_{\lambda \in [0,1]} \{[\rho_{12}^2 + \rho_{21}^2 + 2\rho_{12}\rho_{21}\lambda + 2\delta(\rho_{12}\rho_{21})^2(1-\lambda^2)] / [2\delta - 4\rho_{12}\rho_{21}\lambda - 2\delta(\rho_{12}\rho_{21})^2(1-\lambda^2)]\}. \quad (6.8)$$

Fig. 2 compares the semiexact worst-case performance (6.8) to the majorant Lyapunov equation bound (6.7). To efficiently illustrate the results the data are specialized to the case $\rho_{12} = \rho_{21}$. Note that the semiexact performance is plotted for several values of δ because of the explicit dependence of (6.8) on δ via δ .

ACKNOWLEDGMENT

The authors wish to thank J. Strachla for excellent typing of the original manuscript and B. Bland for executing the numerical calculations.

REFERENCES

- [1] M. G. Safonov and M. Athans, "Gain and phase margin for multiloop LQG regulators," *IEEE Trans. Automat. Contr.*, vol. AC-22, pp. 173-179, 1977.
- [2] M. G. Safonov, *Stability and Robustness of Multivariable Feedback Systems*. Cambridge, MA: M.I.T. Press, 1980.
- [3] N. A. Lehtomaki, N. R. Sandell, Jr., and M. Athans, "Robustness results in linear-quadratic Gaussian based multivariable control designs," *IEEE Trans. Automat. Contr.*, vol. AC-26, pp. 75-92, 1981.
- [4] J. C. Doyle and G. Stein, "Multivariable feedback design: Concepts for a classical/modern synthesis," *IEEE Trans. Automat. Contr.*, vol. AC-26, pp. 4-16, 1981.
- [5] J. C. Doyle, "Analysis of feedback systems with structured uncertainties," *IEEE Proc.*, vol. 129, pp. 242-250, 1982.
- [6] J. C. Doyle, J. E. Wall, and G. Stein, "Performance and robustness analysis for structured uncertainty," in *Proc. 21st IEEE Conf. Decision Contr.*, Orlando, FL, Dec. 1982, pp. 629-636.
- [7] G. Zames, "Feedback and optimal sensitivity: Model reference transformations, multiplicative seminorms, and approximate inverses," *IEEE Trans. Automat. Contr.*, vol. AC-26, pp. 301-320, 1981.
- [8] G. Zames and B. A. Francis, "Feedback, minimax sensitivity, and optimal robustness," *IEEE Trans. Automat. Contr.*, vol. AC-28, pp. 585-601, 1983.
- [9] R. R. E. de Gaston and M. G. Safonov, "A homotopy method for nonconservative stability robustness analysis," in *Proc. 24th IEEE Conf. Decision Contr.*, Fort Lauderdale, FL, Dec. 1985, pp. 1294-1301.
- [10] M. K. H. Fan and A. L. Tits, "Characterization and efficient computation of the structured singular value," *IEEE Trans. Automat. Contr.*, vol. AC-31, pp. 734-743, 1986.
- [11] O. D. I. Nwokah, "The quantitative design of robust multivariable control systems," in *Proc. IEEE Conf. Decision Contr.*, Athens, Greece, Dec. 1986, pp. 16-24.
- [12] I. R. Petersen and C. V. Hollot, "A Riccati equation approach to the stabilization of uncertain systems," *Automatica*, vol. 22, pp. 397-411, 1986.
- [13] R. K. Yedavalli, S. S. Banda, and D. B. Ridgely, "Time-domain stability robustness measures for linear regulators," *J. Guidance Contr. Dyn.*, vol. 8, pp. 520-524, 1985.
- [14] R. K. Yedavalli, "Perturbation bounds for robust stability in linear state space models," *Int. J. Contr.*, vol. 42, pp. 1507-1517, 1985.
- [15] —, "Improved measures of stability robustness for linear state space models," *IEEE Trans. Automat. Contr.*, vol. AC-30, pp. 577-579, 1985.
- [16] A. R. Galimidi and B. R. Barmish, "The constrained Lyapunov problem and its application to robust output feedback stabilization," *IEEE Trans. Automat. Contr.*, vol. AC-31, pp. 410-419, 1986.
- [17] D. S. Bernstein and S. W. Greeley, "Robust output-feedback stabilization: Deterministic and stochastic perspectives," in *Proc. Amer. Contr. Conf.*, Seattle, WA, June 1986, pp. 1818-1826.
- [18] D. S. Bernstein, "Robust static and dynamic output-feedback stabilization: Deterministic and stochastic perspectives," *IEEE Trans. Automat. Contr.*, to be published.
- [19] D. S. Bernstein and D. C. Hyland, "Optimal projection for uncertain systems (OPUS): A unified theory of reduced-order, robust control design," in *Large Space Structures: Dynamics and Control*, S. N. Atluri and A. K. Amos, Eds. New York: Springer-Verlag, 1987.
- [20] W. M. Haddad, "Robust optimal projection control-system synthesis," Ph.D. dissertation, Dep. Mechanical Eng., Florida Inst. Technol., Melbourne, FL, Mar. 1987.
- [21] A. M. Ostrowski, "On some metrical properties of operator matrices

- and matrices partitioned into blocks," *J. Math. Anal. Appl.*, vol. 2, pp. 161-209, 1961.
- [22] T. Strom, "On the practical application of majorants for nonlinear matrix iterations," *J. Math. Anal. Appl.*, vol. 41, pp. 137-147, 1973.
- [23] G. Dahlquist, "On matrix majorants and minorants, with applications to differential equations," *Lin. Alg. Appl.*, vol. 52/53, pp. 199-216, 1983.
- [24] S. Barnett and C. Storey, *Matrix Methods in Stability Theory*. New York: Barnes and Noble, 1976.
- [25] J. W. Brewer, "Kronecker products and matrix calculus in system theory," *IEEE Trans. Circuits Syst.*, vol. CAS-25, pp. 772-781, 1978.
- [26] A. Graham, *Kronecker Products and Matrix Calculus*. Chichester: Ellis Horwood, 1981.
- [27] C. R. Rao and S. K. Mitra, *Generalized Inverse of Matrices and Its Applications*. New York: Wiley, 1971.
- [28] G. P. H. Styan, "Hadamard products and multivariate statistical analysis," *Lin. Alg. Appl.*, vol. 6, pp. 217-240, 1973.
- [29] A. Berman and R. J. Plemmons, *Nonnegative Matrices in the Mathematical Sciences*. New York: Academic, 1979.
- [30] D. D. Siljak, *Large-Scale Dynamic Systems*. Amsterdam, The Netherlands: Elsevier/North-Holland, 1978.
- [31] M. Ikeda and D. D. Siljak, "Generalized decomposition of dynamic systems and vector Lyapunov functions," *IEEE Trans. Automat. Contr.*, vol. AC-26, pp. 1118-1125, 1981.
- [32] M. Djordjevic, "Stability analysis of interconnected systems with possibly unstable subsystems," *Syst. Contr. Lett.*, vol. 3, pp. 165-169, 1983.
- [33] A. A. Martynyuk, "The Lyapunov matrix-function," *Nonlinear Anal. Theory, Methods, and Appl.*, vol. 8, pp. 1223-1226, 1984.
- [34] G. H. Golub and C. F. Van Loan, *Matrix Computations*. Baltimore, MD: Johns Hopkins Press, 1983.
- [35] T. Kato, *Perturbation Theory for Linear Operators*. New York: Springer-Verlag, 1980.
- [36] R. K. Mehra, "Optimization of measurement schedules and sensor designs for linear dynamic systems," *IEEE Trans. Automat. Contr.*, vol. AC-21, pp. 55-64, 1976.
- [37] W. M. Wonham, *Linear Multivariable Control: A Geometric Approach*. New York: Springer-Verlag, 1974.
- [38] G. J. Butler, C. R. Johnson, and H. Wolkowicz, "Nonnegative solutions of a quadratic matrix equation arising from comparison theorems in ordinary differential equations," *SIAM J. Alg. Disc. Meth.*, vol. 6, pp. 47-53, 1985.
- [39] H. D. Victory, Jr., "On nonnegative solutions of matrix equations," *SIAM J. Alg. Discrete Meth.*, vol. 6, pp. 406-412, 1985.



David C. Hyland received the B.S., M.S., and Sc.D. degrees in aeronautics from the Massachusetts Institute of Technology, Cambridge, in 1969, 1971, and 1973, respectively.

After serving as a vibration specialist in a Cambridge-based acoustics consulting firm, in 1974 he joined the staff at Lincoln Laboratory, Massachusetts Institute of Technology, Cambridge. His work at Lincoln Laboratory included reentry vehicle dynamics, multibody spacecraft dynamics simulation, and spacecraft attitude control. In 1983 he joined the Government Aerospace Systems Division, Harris Corporation, Melbourne, FL, where he leads the Control Systems Analysis and Synthesis Group. His current research interests include robustness analysis for control-system design with application to vibration suppression in large flexible space structures.



Dennis S. Bernstein (M'82) received the Sc.B. degree in applied mathematics from Brown University, Providence, RI, and the M.S.E. and Ph.D. degrees from the Computer, Information and Control Engineering Program at the University of Michigan, Ann Arbor.

After spending two years at Lincoln Laboratory, Massachusetts Institute of Technology, Lexington, he joined the Controls Analysis and Synthesis Group of the Government Aerospace Systems Division, Harris Corporation, Melbourne, FL. His current research interests include extending the optimal projection control-design approach to a variety of settings including robust, sampled-data, and decentralized control.

APPENDIX H: Robust Synthesis: Linear Bound

- 95. W. M. Haddad and D. S. Bernstein, "Robust, Reduced-Order, Nonstrictly Proper State Estimation via the Optimal Projection Equations with Guaranteed Cost Bounds," IEEE Trans. Autom. Contr., Vol. 33, pp. 591-595, 1988.
- 119. D. S. Bernstein and W. M. Haddad, "Robust Stability and Performance for Fixed-Order Dynamic Compensation via the Optimal Projection Equations with Guaranteed Cost Bounds," Math. Contr. Sig. Sys., to appear.
- 89. W. M. Haddad and D. S. Bernstein, "The Unified Optimal Projection Equations for Simultaneous Reduced-Order, Robust Modeling, Estimation and Control," Int. J. Contr., Vol. 47, pp. 1117-1132, 1988.

Robust, Reduced-Order, Nonstrictly Proper State Estimation Via the Optimal Projection Equations with Guaranteed Cost Bounds

WASSIM M. HADDAD AND DENNIS S. BERNSTEIN

Abstract—A state-estimation design problem involving parametric plant uncertainties is considered. An estimation error bound suggested by multiplicative white noise modeling is utilized for guaranteeing robust estimation over a specified range of parameter uncertainties. Necessary conditions which generalize the optimal projection equations for reduced-order state estimation are used to characterize the estimator which minimizes the error bound. The design equations thus effectively serve as sufficient conditions for synthesizing robust estimators. Additional features include the presence of a static estimation gain in conjunction with the dynamic (Kalman) estimator to obtain a nonstrictly proper estimator.

I. INTRODUCTION

As is well known [1]–[12], the performance of optimal filters based upon nominal parameter values may be severely degraded in the presence of parameter deviations. Thus, it is desirable to obtain *robust* state estimators which provide acceptable performance over the range of parametric uncertainty. The approach of the present paper is related to the guaranteed cost approach developed for control in [13], [14] and applied to estimation in [3]. Specifically, the main idea is to bound the effect of the uncertain parameters on the estimation error over the uncertainty range and then choose estimator gains to minimize the estimation bound. Thus, the *actual* estimation error is guaranteed to lie below the prescribed upper bound.

The technique used to determine minimizing estimator gains is a generalization of the optimal projection equations for reduced-order state estimation [15]. Thus, the results of the present paper effectively extend the results of [15] to the case of parameter uncertainties. It should be noted that the optimal projection equations, which are necessary conditions for optimality, now serve as *sufficient* conditions for robust estimation by virtue of the fact that a bound on the estimation error is being minimized rather than the estimation error itself. The bound utilized in the present paper was originally suggested by multiplicative white noise

Manuscript received June 8, 1987; revised September 25, 1987. This work was supported in part by the U.S. Air Force Office of Scientific Research under Contract F49620-86-C-0002.

W. M. Haddad is with the Department of Mechanical Engineering, Florida Institute of Technology, Melbourne, FL 32901.

D. S. Bernstein is with the Government Aerospace Systems Division, Harris Corporation, Melbourne, FL 32902.

IEEE Log Number 8718565.

modeling and was used in [16]–[18] for constructing Lyapunov functions for robust fixed-order dynamic compensation. A similar bound was used for full-state feedback in [19].

An additional feature of the present paper is the inclusion of a static feedback gain in conjunction with the dynamic estimator. Thus, the results of the present paper represent a generalization of standard results to the case of nonstrictly proper estimation. Similar treatments in the context of multiplicative noise models were given in [10] and [11] for discrete-time and continuous-time systems, respectively.

II. NOTATION AND DEFINITIONS

Note: All matrices have real entries.

| | |
|---|--|
| $\mathbb{R}, \mathbb{R}^{r \times s}, \mathbb{R}^r$ | Real numbers, $r \times s$ real matrices, $\mathbb{R}^{r \times 1}$. |
| $I_r, (\cdot)^T, \mathbb{E}$ | $r \times r$ identity matrix, transpose, expected value. |
| \oplus, \otimes | Kronecker sum, Kronecker product [20]. |
| $\mathbb{S}', \mathbb{N}', \mathbb{P}'$ | $r \times r$ symmetric, nonnegative-definite, positive-definite matrices. |
| $Z_1 \leq Z_2, Z_1 < Z_2$ | $Z_2 - Z_1 \in \mathbb{N}', Z_2 - Z_1 \in \mathbb{P}', Z_1, Z_2 \in \mathbb{S}'$. |
| $n, l, \bar{l}, n_e, p, q; \bar{n}$ | Positive integers; $n + n_e$. |
| $x, y, \hat{x}, y_e, x_e, \hat{x}$ | $n, l, \bar{l}, q, n_e, \bar{n}$ -dimensional vectors. |
| $A, \Delta A; C, \Delta C$ | $n \times n$ matrices; $l \times n$ matrices. |
| \hat{C} | $\bar{l} \times n$ matrix. |
| A_e, B_e, C_e, D_e | $n_e \times n_e, n_e \times l, q \times n_e, q \times \bar{l}$ matrices. |
| $\bar{A}, \Delta \bar{A}$ | $\begin{bmatrix} A & 0 \\ B_e C & A_e \end{bmatrix}, \begin{bmatrix} \Delta A & 0 \\ B_e \Delta C & 0 \end{bmatrix}$. |
| L, R | $q \times n$ matrix, estimation-error weighting in \mathbb{P}^q . |
| \bar{R} | $\begin{bmatrix} L^T R L - L^T R D_e \hat{C} - \hat{C}^T D_e^T R L + \hat{C}^T D_e^T R D_e \hat{C} & -L^T R C_e + \hat{C}^T D_e^T R C_e \\ -\hat{C}^T R L + \hat{C}^T R D_e \hat{C} & C_e^T R C_e \end{bmatrix}$. |
| $w_1(\cdot), w_2(\cdot)$ | n, l -dimensional white noise. |
| V_1, V_2 | Intensity of $w_1(\cdot), w_2(\cdot)$; $V_1 \in \mathbb{N}', V_2 \in \mathbb{P}'$. |
| V_{12} | Cross intensity of $w_1(\cdot), w_2(\cdot)$. |
| $\bar{w}(\cdot), \bar{V}$ | $\begin{bmatrix} w_1(\cdot) \\ B_e w_2(\cdot) \end{bmatrix}, \begin{bmatrix} V_1 & V_{12} B_e^T \\ B_e V_{12}^T & B_e V_2 B_e^T \end{bmatrix}$. |
| α | Positive number. |
| α_i | Positive number, $i = 1, \dots, p$. |
| σ_i | Real number, $i = 1, \dots, p$. |
| γ_i | $\alpha_i^2 / \alpha, i = 1, \dots, p$. |
| $A_\alpha, A_{e\alpha}$ | $A + (\alpha/2)I_n, A_e + (\alpha/2)I_{n_e}$. |
| \bar{A}_α | $\begin{bmatrix} A_\alpha & 0 \\ B_e C & A_{e\alpha} \end{bmatrix}$. |

III. ROBUST ESTIMATION PROBLEM

Let $\mathcal{U} \subset \mathbb{R}^{n \times n} \times \mathbb{R}^{l \times n}$ denote the set of uncertain perturbations $(\Delta A, \Delta C)$ of the nominal plant matrices A and C .

Robust Estimation Problem: For fixed $n_e \leq n$, determine (A_e, B_e, C_e, D_e) such that, for the system consisting of the n th-order disturbed plant

$$\dot{x}(t) = (A + \Delta A)x(t) + w_1(t), \quad t \in [0, \infty), \quad (3.1)$$

noisy and nonnoisy measurements

$$y(t) = (C + \Delta C)x(t) + w_2(t), \quad (3.2)$$

$$\hat{y}(t) = \hat{C}x(t), \quad (3.3)$$

and n_e th-order nonstrictly proper state estimator

$$\dot{x}_e(t) = A_e x_e(t) + B_e y(t), \quad (3.4)$$

$$y_e(t) = C_e x_e(t) + D_e \hat{y}(t), \quad (3.5)$$

the state-estimation error criterion

$$J(A_e, B_e, C_e, D_e) \triangleq \sup_{(\Delta A, \Delta C) \in \mathcal{U}} \limsup_{t \rightarrow \infty} \mathbb{E} \{ [x(t) - y_e(t)]^T R [x(t) - y_e(t)] \} \quad (3.6)$$

is minimized.

For each estimator (A_e, B_e, C_e, D_e) and system variation $(\Delta A, \Delta C) \in \mathcal{U}$, the disturbed augmented system (3.1)–(3.5) is given by

$$\dot{\bar{x}}(t) = (\bar{A} + \Delta \bar{A})\bar{x}(t) + \bar{w}(t), \quad t \in [0, \infty) \quad (3.7)$$

where $\bar{x}(t) \triangleq [x^T(t), x_e^T(t)]^T$ and $\bar{w}(t)$ has intensity $\bar{V} \in \mathbb{N}^{\bar{n}}$. The cost can be expressed in terms of the second-moment matrix.

Proposition 3.1: For given (A_e, B_e, C_e, D_e) and $(\Delta A, \Delta C) \in \mathcal{U}$, the second-moment matrix

$$\bar{Q}_{\Delta \bar{A}}(t) \triangleq \mathbb{E} \{ \bar{x}(t) \bar{x}^T(t) \}, \quad t \in [0, \infty) \quad (3.8)$$

satisfies

$$\dot{\bar{Q}}_{\Delta \bar{A}}(t) = (\bar{A} + \Delta \bar{A})\bar{Q}_{\Delta \bar{A}}(t) + \bar{Q}_{\Delta \bar{A}}(t)(\bar{A} + \Delta \bar{A})^T + \bar{V}, \quad t \in [0, \infty). \quad (3.9)$$

Furthermore,

$$J(A_e, B_e, C_e, D_e) \triangleq \sup_{(\Delta A, \Delta C) \in \mathcal{U}} \limsup_{t \rightarrow \infty} \text{tr} \bar{Q}_{\Delta \bar{A}}(t) \bar{R}. \quad (3.10)$$

IV. SUFFICIENT CONDITIONS FOR ROBUST PERFORMANCE

The following result is immediate.

Lemma 4.1: Suppose $\bar{A} + \Delta \bar{A}$ is stable for all $(\Delta A, \Delta C) \in \mathcal{U}$. Then

$$J(A_e, B_e, C_e, D_e) = \sup_{(\Delta A, \Delta C) \in \mathcal{U}} \text{tr} \bar{Q}_{\Delta \bar{A}} \bar{R} \quad (4.1)$$

where $\bar{Q}_{\Delta A} \in \mathbb{R}^n$ is the unique solution to

$$0 = (\bar{A} + \Delta \bar{A})\bar{Q}_{\Delta A} + \bar{Q}_{\Delta A}(\bar{A} + \Delta \bar{A})^T + \bar{P}. \quad (4.2)$$

We seek upper bounds for $J(A_e, B_e, C_e, D_e)$.

Theorem 4.1: Let $\Omega: \mathbb{R}^n \times \mathbb{R}^{n \times n} \rightarrow \mathbb{S}^n$ be such that

$(\Delta A, \Delta C) \in \mathcal{U}$,

$$\Delta \bar{A}Q + Q\Delta \bar{A}^T \leq \Omega(Q, B_e), (Q, B_e) \in \mathbb{R}^n \times \mathbb{R}^{n \times n} \quad (4.3)$$

and, for given (A_e, B_e, C_e, D_e) , suppose there exists $Q \in \mathbb{R}^n$ satisfying

$$0 = \bar{A}Q + Q\bar{A}^T + \Omega(Q, B_e) + \bar{P}, \quad (4.4)$$

and suppose the pair $(\bar{P}^{1/2}, \bar{A} + \Delta \bar{A})$ is stabilizable for all $(\Delta A, \Delta C) \in \mathcal{U}$. Then A_e is asymptotically stable, $A + \Delta A$ is asymptotically stable for all $(\Delta A, \Delta C) \in \mathcal{U}$,

$$\bar{Q}_{\Delta A} \leq Q, \quad (\Delta A, \Delta C) \in \mathcal{U} \quad (4.5)$$

where $\bar{Q}_{\Delta A}$ satisfies (4.2) and

$$J(A_e, B_e, C_e, D_e) \leq \text{tr } Q\bar{R}. \quad (4.6)$$

Proof: For all $(\Delta A, \Delta C) \in \mathcal{U}$, (4.4) is equivalent to

$$0 = (\bar{A} + \Delta \bar{A})Q + Q(\bar{A} + \Delta \bar{A})^T + \Psi(Q, B_e, \Delta \bar{A}) + \bar{P} \quad (4.7)$$

where

$$\Psi(Q, B_e, \Delta \bar{A}) \triangleq \Omega(Q, B_e) - (\Delta \bar{A}Q + Q\Delta \bar{A}^T).$$

Note that by (4.3), $\Psi(Q, B_e, \Delta \bar{A}) \geq 0$ for all $(\Delta A, \Delta C) \in \mathcal{U}$. Since $(\bar{P}^{1/2}, \bar{A} + \Delta \bar{A})$ is stabilizable for all $(\Delta A, \Delta C) \in \mathcal{U}$, it follows from [21, Theorem 3.6] that $((\bar{P} + \Psi(Q, B_e, \Delta \bar{A}))^{1/2}, \bar{A} + \Delta \bar{A})$ is stabilizable for all $(\Delta A, \Delta C) \in \mathcal{U}$. Hence, [21, Lemma 12.2] implies $\bar{A} + \Delta \bar{A}$ is asymptotically stable for all $(\Delta A, \Delta C) \in \mathcal{U}$. Since $\bar{A} + \Delta \bar{A}$ is lower block triangular, A_e is asymptotically stable and $A + \Delta A$ is asymptotically stable for all $(\Delta A, \Delta C) \in \mathcal{U}$. Next, (4.7) minus (4.2) yields

$$0 = (\bar{A} + \Delta \bar{A})(Q - \bar{Q}_{\Delta A}) + (Q - \bar{Q}_{\Delta A})(\bar{A} + \Delta \bar{A})^T + \Psi(Q, B_e, \Delta \bar{A})$$

or, equivalently (since $\bar{A} + \Delta \bar{A}$ is asymptotically stable),

$$Q - \bar{Q}_{\Delta A} = \int_0^\infty e^{(\bar{A} + \Delta \bar{A})t} \Psi(Q, B_e, \Delta \bar{A}) e^{(\bar{A} + \Delta \bar{A})^T t} dt \geq 0,$$

which implies (4.5). Finally, (4.5) and (4.1) yield (4.6). \square

V. UNCERTAINTY STRUCTURE AND GUARANTEED COST BOUND

The uncertainty set \mathcal{U} is assumed to be of the form

$$\mathcal{U} \triangleq \left\{ (\Delta A, \Delta C) \in \mathbb{R}^{n \times n} \times \mathbb{R}^{l \times n} : \Delta A = \sum_{i=1}^p \sigma_i A_i, \Delta C = \sum_{i=1}^p \sigma_i C_i, \sum_{i=1}^p \sigma_i^2 / \alpha_i^2 \leq 1 \right\} \quad (5.1)$$

where, for $i = 1, \dots, p$, $A_i \in \mathbb{R}^{n \times n}$ and $C_i \in \mathbb{R}^{l \times n}$ are fixed matrices denoting the structure of the parametric uncertainty in the dynamics and measurement matrices; α_i is a given positive number; and σ_i is an uncertain real parameter. In practice, the form of ΔA and ΔC permits the modeling of linear parameter uncertainties of arbitrary structure. Note that the uncertain parameters σ_i are assumed to lie in a specified ellipsoidal region in \mathbb{R}^p . The augmented system thus has structured uncertainty of the form

$$\Delta \bar{A} = \sum_{i=1}^p \sigma_i \bar{A}_i \quad (5.2)$$

where

$$\bar{A}_i \triangleq \begin{bmatrix} A_i & 0 \\ B_e C_i & 0 \end{bmatrix}, \quad i = 1, \dots, p.$$

Remark 5.1: Note that (5.1) allows a particular parameter σ_i to appear in both ΔA and ΔC . Thus, it is possible to consider the case in which the uncertainties ΔA and ΔC are known to be correlated. Of course, for a given i , A_i or C_i can be set to zero so that the similar form of ΔA and ΔC represents no restriction.

We now specify the bounding function Ω satisfying (4.3).

Proposition 5.1: Let α be an arbitrary positive scalar. Then the function

$$\Omega(Q, B_e) \triangleq \alpha Q + \alpha^{-1} \sum_{i=1}^p \alpha_i^2 \bar{A}_i Q \bar{A}_i^T \quad (5.3)$$

satisfies (4.3) with \mathcal{U} given by (5.1).

Proof: Note that

$$\begin{aligned} 0 &\leq \sum_{i=1}^p [(\alpha^{1/2} \sigma_i / \alpha_i) I_n - (\alpha_i / \alpha^{1/2}) \bar{A}_i] \cdot Q [(\alpha^{1/2} \sigma_i / \alpha_i) I_n - (\alpha_i / \alpha^{1/2}) \bar{A}_i]^T \\ &= \alpha \sum_{i=1}^p (\sigma_i^2 / \alpha_i^2) Q + \alpha^{-1} \sum_{i=1}^p \alpha_i^2 \bar{A}_i Q \bar{A}_i^T - \sum_{i=1}^p \sigma_i (\bar{A}_i Q + Q \bar{A}_i^T) \end{aligned}$$

which, since $\sum_{i=1}^p \sigma_i^2 / \alpha_i^2 \leq 1$, implies (4.3). \square

Remark 5.2: Note that with (5.3), the modified Lyapunov equation (4.4) becomes

$$0 = \bar{A}_e Q + Q \bar{A}_e^T + \sum_{i=1}^p \gamma_i \bar{A}_i Q \bar{A}_i^T + \bar{P}. \quad (5.4)$$

VI. THE AUXILIARY MINIMIZATION PROBLEM

Our goal is to minimize the error bound (4.6).

Auxiliary Minimization Problem: Determine (Q, A_e, B_e, C_e, D_e) with $Q \in \mathbb{R}^n$ which minimizes

$$\mathcal{J}(Q, A_e, B_e, C_e, D_e) \triangleq \text{tr } Q\bar{R} \quad (6.1)$$

subject to (5.4) and

$$(\bar{P}^{1/2}, \bar{A} + \Delta \bar{A}) \text{ is stabilizable}, \quad (\Delta A, \Delta C) \in \mathcal{U}. \quad (6.2)$$

Proposition 6.1: If (Q, A_e, B_e, C_e, D_e) satisfies (5.4) and (6.2) with $Q \geq 0$, then $\bar{A} + \Delta \bar{A}$ is asymptotically stable for all $(\Delta A, \Delta C) \in \mathcal{U}$ and

$$J(A_e, B_e, C_e, D_e) \leq \mathcal{J}(Q, A_e, B_e, C_e, D_e). \quad (6.3)$$

Proof: With Ω given by (5.3), Proposition 5.1 implies that (4.3) is satisfied. Hence, with (6.2), the hypotheses of Theorem 4.1 are satisfied so that the system (3.7) is stable over \mathcal{U} with estimation bound (4.6). Note that (6.3) is merely a restatement of (4.6). \square

Remark 6.1: The conservatism of the bound (6.3) is difficult to predict for two reasons. First, the overbounding (4.3) holds with respect to the partial ordering of the nonnegative-definite matrices for which no scalar measure of conservatism is available. And second, the bound (4.3) is required to hold for all nonnegative-definite matrices Q and estimator gains B_e . The conservatism will thus depend upon the actual values of Q and B_e determined by solving (5.4).

VII. NECESSARY CONDITIONS FOR THE AUXILIARY MINIMIZATION PROBLEM

Rigorous application of the Lagrange multiplier technique requires additional technical assumptions. Specifically, we further restrict $(Q, A_e,$

B_e, C_e, D_e) to the open set

$$\mathcal{S} \triangleq \{(Q, A_e, B_e, C_e, D_e) : Q \in \mathbb{P}^n,$$

\tilde{Q} is asymptotically stable,

(A_e, B_e, C_e) is controllable and observable, and

$$\tilde{C}(Q_1 - Q_{12}Q_2^{-1}Q_{12}^T)\tilde{C}^T > 0\}$$

where

$$\tilde{Q} \triangleq \tilde{A}_e \oplus \tilde{A}_e + \sum_{i=1}^p \gamma_i \tilde{A}_i \oplus \tilde{A}_i$$

and Q is partitioned as

$$Q \triangleq \begin{bmatrix} Q_1 & Q_{12} \\ Q_{12}^T & Q_2 \end{bmatrix}$$

where Q_1, Q_{12} , and Q_2 are $n \times n$, $n \times n_e$, and $n_e \times n_e$, respectively. As shown in [11], Q_2 is invertible since (A_e, B_e) is controllable. The definiteness condition holds when \tilde{C} has full row rank and Q is positive definite. As shown in [11], this condition implies the existence of the projection $\hat{\tau}$ defined below.

Remark 7.1: Proposition 6.1 shows that the constraint $(Q, A_e, B_e, C_e, D_e) \in \mathcal{S}$ is not required for robust estimation. As can be seen from the proof given in [11], the set \mathcal{S} constitutes sufficient conditions under which the Lagrange multiplier technique is applicable to the auxiliary minimization problem. Specifically, $Q \in \mathbb{P}^n$ replaces $Q \in \mathbb{N}^n$ by an open set constraint, while asymptotic stability of \tilde{Q} serves as a normality condition which further implies that the dual \mathcal{Q} of Q is nonnegative definite. Thus, it is not necessary for guaranteed robust estimation that an admissible quadruple obtained by solving the necessary conditions actually be shown to be an element of \mathcal{S} .

The following factorization lemma is needed for the statement of the main result. For details, see [15].

Lemma 7.1: If $\tilde{Q}, \tilde{P} \in \mathbb{N}^n$ and $\text{rank } \tilde{Q}\tilde{P} = n_e$, then there exist $n_e \times n$ G, Γ , and $n_e \times n_e$ invertible M such that

$$\tilde{Q}\tilde{P} = G^T M \Gamma \quad (7.1)$$

$$\Gamma G^T = I_{n_e}. \quad (7.2)$$

Recall from [15] that

$$\tau \triangleq \tilde{Q}\tilde{P}(\tilde{Q}\tilde{P})^\dagger = G^T \Gamma \quad (7.3)$$

is an oblique projection. Define the complementary projection $\tau_\perp \triangleq I_n - \tau$ and call (G, M, Γ) satisfying (7.1), (7.2) a *projective factorization* of $\tilde{Q}\tilde{P}$. Furthermore, for arbitrary $Q, \tilde{Q} \in \mathbb{N}^{n \times n}$, define the notation

$$V_{2i} \triangleq V_2 + \sum_{i=1}^p \gamma_i C_i (Q + \tilde{Q}) C_i^T.$$

$$Q_i \triangleq Q C_i^T + V_{2i} + \sum_{i=1}^p \gamma_i A_i (Q + \tilde{Q}) C_i^T.$$

Theorem 7.1: If $(Q, A_e, B_e, C_e, D_e) \in \mathcal{S}$ solves the auxiliary minimization problem with \mathcal{U} given by (5.1) and Ω given by (5.3), then there exist $Q, \tilde{Q}, \tilde{P} \in \mathbb{N}^n$ such that, for some projective factorization (G, M, Γ) of $\tilde{Q}\tilde{P}$, (Q, A_e, B_e, C_e, D_e) are given by

$$Q = \begin{bmatrix} Q + \tilde{Q} & Q\Gamma^T \\ \Gamma Q & \Gamma Q \Gamma^T \end{bmatrix}, \quad (7.4)$$

$$A_e = \Gamma(A - Q_e V_{2e}^{-1} C) G^T, \quad (7.5)$$

$$B_e = \Gamma Q_e V_{2e}^{-1}, \quad (7.6)$$

$$C_e = L \hat{\tau}_\perp G^T, \quad (7.7)$$

$$D_e = L Q C^T (C Q C^T)^{-1}, \quad (7.8)$$

and such that Q, \tilde{Q} , and \tilde{P} satisfy

$$0 = A_e Q + Q A_e^T + V_1 + \sum_{i=1}^p \gamma_i A_i (Q + \tilde{Q}) A_i^T - Q_e V_{2e}^{-1} Q_e^T + \tau_\perp Q_e V_{2e}^{-1} Q_e^T \tau_\perp^T, \quad (7.9)$$

$$0 = A_e \tilde{Q} + \tilde{Q} A_e^T + Q_e V_{2e}^{-1} Q_e^T - \tau_\perp Q_e V_{2e}^{-1} Q_e^T \tau_\perp^T, \quad (7.10)$$

$$0 = (A_e - Q_e V_{2e}^{-1} C)^T \tilde{P} + \tilde{P} (A_e - Q_e V_{2e}^{-1} C) + \hat{\tau}_\perp^T L^T R L \hat{\tau}_\perp - \tau_\perp^T \hat{\tau}_\perp^T L^T R L \tau_\perp, \quad (7.11)$$

$$\text{rank } \tilde{Q} = \text{rank } \tilde{P} = \text{rank } \tilde{Q}\tilde{P} = n_e, \quad (7.12)$$

where

$$\hat{\tau} \triangleq Q C^T (C Q C^T)^{-1} C, \quad \hat{\tau}_\perp \triangleq I_n - \hat{\tau}. \quad (7.13)$$

Furthermore, the auxiliary cost is given by

$$J(Q, A_e, B_e, C_e, D_e) = \text{tr } Q \hat{\tau}_\perp^T L^T R L \hat{\tau}_\perp. \quad (7.14)$$

Conversely, if there exist $Q, \tilde{Q}, \tilde{P} \in \mathbb{N}^n$ satisfying (7.9)–(7.12), then (Q, A_e, B_e, C_e, D_e) given by (7.4)–(7.8) satisfy (5.4) with $Q \in \mathbb{N}^n$ and with $\mathcal{J}(Q, A_e, B_e, C_e, D_e)$ given by (7.14).

Proof: The derivation requires only a minor modification of the derivation given in [11]. The only change involves treatment of \tilde{A}_e in place of \tilde{A} . \square

Remark 7.1: The necessary conditions given in Theorem 7.1 directly generalize the result given in [15]. To recover the result of [15], set $A_i = 0, C_i = 0, i = 1, \dots, p$ (to delete the plant uncertainties), and set $\tilde{C} = 0$ (to eliminate the static estimation term D_e). It follows from the proof given in [11] that $\tilde{C} = 0$ yields $\hat{\tau} = 0$, and thus $\hat{\tau}_\perp = I_n$.

Remark 7.2: Note that Q given by (7.4) is nonnegative definite.

VIII. SUFFICIENT CONDITIONS FOR ROBUST, REDUCED-ORDER ESTIMATION

The main result guaranteeing robust estimation can now be stated.

Theorem 8.1: Suppose there exist $Q, \tilde{Q}, \tilde{P} \in \mathbb{N}^n$ satisfying (7.9)–(7.12), let A_e, B_e, C_e, D_e be given by (7.5)–(7.8), and suppose that $(\tilde{P}^{1/2}, \tilde{A} + \Delta \tilde{A})$ is stabilizable for all $(\Delta A, \Delta C) \in \mathcal{U}$ with \mathcal{U} given by (5.1). Then A_e is asymptotically stable, $A + \Delta A$ is asymptotically stable for all $(\Delta A, \Delta C) \in \mathcal{U}$, and the estimation error satisfies the performance bound

$$J(A_e, B_e, C_e, D_e) \leq \text{tr } Q \hat{\tau}_\perp^T L^T R L \hat{\tau}_\perp. \quad (8.1)$$

Proof: Theorem 7.1 and Remark 7.2 imply that Q given by (7.4) is nonnegative definite and satisfies (5.4). With the stabilizability assumption, the result follows from Proposition 6.1. \square

Remark 8.1: Suppose $\hat{f} = n, \hat{C} = I_n$ (so that perfect measurements of the entire state are available), and Q satisfying (7.9) is positive definite. Then it follows from Theorem 7.1 that $\hat{\tau} = I_n, \hat{\tau}_\perp = 0, C_e = 0$ (i.e., the dynamic filter is disabled), $D_e = L$, and by (8.1), $J = 0$. This is, of course, the expected result since perfect estimation is achievable in this case.

REFERENCES

- [1] J. A. D'Appolito and C. E. Hutchinson, "Low sensitivity filters for state estimation in the presence of large parameter uncertainties," *IEEE Trans. Automat. Contr.*, vol. AC-14, pp. 310–312, 1969.
- [2] P. J. McLane, "Optimal linear filtering for linear systems with state-dependent noise," *Int. J. Contr.*, vol. 10, pp. 41–51, 1969.
- [3] B. N. Jain, "Guaranteed error estimation in uncertain systems," *IEEE Trans. Automat. Contr.*, vol. AC-20, pp. 230–232, 1975.

- [4] R. C. Chung and P. R. Belanger, "Minimum-sensitivity filter for linear time-invariant stochastic systems with uncertain parameters," *IEEE Trans. Automat. Contr.*, vol. AC-21, pp. 98-100, 1976.
- [5] K. Furuta, S. Hara, and S. Mori, "A class of systems with the same observer," *IEEE Trans. Automat. Contr.*, vol. AC-21, pp. 572-576, 1976.
- [6] S. P. Bhattacharyya, "The structure of robust observers," *IEEE Trans. Automat. Contr.*, vol. AC-21, pp. 581-588, 1976.
- [7] M. Toda and R. V. Patel, "Bounds on estimation errors of discrete-time filters under modeling uncertainty," *IEEE Trans. Automat. Contr.*, vol. AC-25, pp. 1115-1121, 1980.
- [8] R. T. Stefani, "Reducing the sensitivity to parameter variations of a minimum-order reduced-order observer," *Int. J. Contr.*, vol. 35, pp. 983-995, 1982.
- [9] T. E. Djaferis, "Robust observers for systems with parameters," *Syst. Contr. Lett.*, vol. 7, pp. 385-394, 1986.
- [10] W. M. Haddad and D. S. Bernstein, "The optimal projection equations for discrete-time reduced-order state estimation for linear systems with multiplicative white noise," *Syst. Contr. Lett.*, vol. 8, pp. 381-388, 1987.
- [11] —, "The optimal projection equations for reduced-order state estimation: The singular measurement noise case," *IEEE Trans. Automat. Contr.*, vol. AC-32, pp. 1135-1139, 1987.
- [12] D. S. Bernstein and D. C. Hyland, "The optimal projection equations for reduced-order modelling, estimation and control of linear systems with multiplicative white noise," *J. Optimiz. Theory Appl.*, to be published.
- [13] S. S. L. Chang and T. K. C. Peng, "Adaptive guaranteed cost control of systems with uncertain parameters," *IEEE Trans. Automat. Contr.*, vol. AC-17, pp. 474-483, 1972.
- [14] A. Vinkler and L. J. Wood, "Multistep guaranteed cost control of linear systems with uncertain parameters," *J. Guidance Contr.*, vol. 2, pp. 449-456, 1979.
- [15] D. S. Bernstein and D. C. Hyland, "The optimal projection equations for reduced-order state estimation," *IEEE Trans. Automat. Contr.*, vol. AC-30, pp. 583-585, 1985.
- [16] D. S. Bernstein and S. W. Greeley, "Robust output-feedback stabilization: Deterministic and stochastic perspectives," in *Proc. Amer. Contr. Conf.*, Seattle, WA, June 1986, pp. 1818-1826.
- [17] W. M. Haddad, "Robust optimal projection control-system synthesis," Ph.D. dissertation, Dept. Mech. Eng., Florida Inst. Technol., Melbourne, Mar. 1987.
- [18] D. S. Bernstein, "Robust static and dynamic output-feedback stabilization: Deterministic and stochastic perspectives," *IEEE Trans. Automat. Contr.*, vol. AC-32, pp. 1076-1084, 1987.
- [19] O. I. Kosmidou and P. Bertrand, "Robust-controller design for systems with large parameter variations," *Int. J. Contr.*, vol. 45, pp. 927-938, 1987.
- [20] J. W. Brewer, "Kronecker products and matrix calculus in system theory," *IEEE Trans. Circuits Syst.*, vol. CAS-25, pp. 772-781, 1978.
- [21] W. M. Wonham, *Linear Multivariable Control: A Geometric Approach*, New York: Springer-Verlag, 1979.

Approximate and Limit Results for Nonlinear Filters with Small Observation Noise: The Linear Sensor and Constant Diffusion Coefficient Case

OFFER ZEITOUNI

Abstract—Recursive approximations for a class of filtering problems are presented. This class is characterized by linear observation sensor, constant diffusion terms, and for the multidimensional problem, potential-like conditions on the drift. For the case of small observation noise, these approximations are used to demonstrate the Gaussian limiting structure of the optimal nonlinear filter.

I. INTRODUCTION

The classical nonlinear filtering problem is of the form

$$dx_t = f(x_t)dt + \sigma(x_t)dw_t, \quad x_t \in \mathbb{R}^n, \quad p(x_0) = p_0(x_0) \quad (1.1)$$

$$dy_t = g(x_t)dt + N_0^{1/2}d\theta_t, \quad y_t \in \mathbb{R}^m, \quad y_0 = 0 \quad (1.2)$$

Manuscript received February 5, 1987; revised September 2, 1987. Paper recommended by Associate Editor, T. J. Tarn. This work was supported in part by a Weizmann Postdoctoral Fellowship.

The author was with the Technion—Israel Institute of Technology, Haifa 32000, Israel. He is now with LIDS, Massachusetts Institute of Technology, Cambridge, MA 02139.

IEEE Log Number 8718566.

where w_t, θ_t are independent Brownian motions and the filtering problem consists of computing statistics of x_t when the observation σ algebra $y_0^t \triangleq \{y_s, 0 \leq s \leq t\}$ is given. By now, it is clear that for all but a few problems, an explicit final dimensional solution does not exist [4]. Therefore, one is led to consider approximations and to consider simplified limiting cases. Especially, the low observation noise case ($N_0 \rightarrow 0$) has been considered in the literature [3], [6], [11].

In this paper, we restrict our attention to the special case of linear observations and constant diffusion coefficients, i.e.,

$$dx_t = f(x_t)dt + \sigma dw_t, \quad x_t \in \mathbb{R}^n, \quad p(x_0) = p_0(x_0) \quad (1.1')$$

$$dy_t = gx_t dt + N_0^{1/2}d\theta_t, \quad y_t \in \mathbb{R}^m, \quad y_0 = 0 \quad (1.2')$$

where σ, g are matrices of appropriate dimensions. In the multidimensional case, we impose some additional potential-like structural conditions on $f(\cdot)$. For this restricted class of filtering problems, we derive recursive approximations to the conditional density $p_{x_t}(z|y_0^t)$ and to its unnormalized version $\rho_{x_t}(z|y_0^t)$. For the limiting case $N_0 \rightarrow 0$, those approximations are used to show that the conditional density, rescaled in a suitable manner, converges to a Gaussian density, with tight estimates on the "tails" of the density. This fact demonstrates the usefulness of Gaussian-based approximations (like the extended Kalman filter or the second-order Gaussian filter).

Related results were obtained by Mayer-Wolf [9] in his dissertation. There, bounds on the filtering error and the Cramer-Rao inequality are used to prove a basic Gaussian limit result, although under different assumptions.

The paper is organized as follows. The one-dimensional problem ($n = m = 1$) is treated in Sections II and III. In Section II, we present our basic approximation theorem, which holds whether N_0 is small or not. We further demonstrate that, if $N_0 \rightarrow 0$ the approximations exhibit certain nice limiting behavior, then the rescaled conditional density converges (weakly and pointwise) to a Gaussian density. In Section III, we check out explicitly the limiting behavior of the approximations and derive explicit conditions on $f(\cdot)$ under which the density indeed converges to a Gaussian one. Finally, in Section IV, we extend our results to a class of multidimensional problems.

We make throughout, the following assumption.

(A1) $f(\cdot)$ is continuously differentiable with bounded first partial derivatives.

II. AN APPROXIMATION THEOREM—THE ONE-DIMENSIONAL CASE

In this section, an approximation theorem for the unnormalized conditional density $\rho(z|y_0^t)$ is presented. Throughout, the one-dimensional case is treated. Multidimensional extensions are postponed to Section IV.

Without loss of generality, we assume $\sigma = 1$ in (1.1'). Recall that under (A-1), a solution to (1.1') exists and is unique. Moreover, the measure P_t defined by the pair (1.1'), (1.2') is absolutely continuous w.r.t. the reference measure P_0 defined by

$$dx_t = \alpha x_t dt + dw_t, \quad p(x_0) = p_0(x_0) \quad (2.1)$$

$$dy_t = N_0^{1/2}d\theta_t, \quad y_0 = 0 \quad (2.2)$$

where α is some constant to be defined. The Radon-Nikodym derivative dP_t/dP_0 is [8].

$$\frac{dP_t}{dP_0} = \exp \left(\int_0^t (f(x_s) - \alpha x_s) dx_s - \frac{1}{2} \int_0^t (f^2(x_s) - \alpha^2 x_s^2) ds + \int_0^t N_0^{-1} g x_s dy_s - \frac{1}{2} \int_0^t N_0^{-1} g^2 x_s^2 ds \right) \quad (2.3)$$

As is well known [2], [12], [13], the unnormalized conditional density $\rho(z|y_0^t)$ satisfies

$$\rho(z|y_0^t) \propto E_0 \left(\frac{dP_t}{dP_0} \middle| y_0^t, x_t = z \right) p_{x_t}(z) \quad (2.4)$$

Submitted July 1987
Revised July 1988

**Robust Stability and Performance via
Fixed-Order Dynamic Compensation
with
Guaranteed Cost Bounds**

by

Dennis S. Bernstein
Harris Corporation
Government Aerospace Systems Division
MS 22/4848
Melbourne, FL 32902

Wassim M. Haddad
Department of Mechanical and
Aerospace Engineering
Florida Institute of Technology
Melbourne, FL 32901

Key words: robust, guaranteed bounds, LQG control, reduced-order, optimal, structured uncertainty, real parameters

Abstract

A feedback control-design problem involving structured real-valued plant parameter uncertainties is considered. Two robust control-design issues are addressed. The Robust Stability Problem involves deterministic bounded structured parameter variations, while the Robust Performance Problem includes, in addition, a quadratic performance criterion averaged over stochastic disturbances and maximized over the admissible parameter variations. The optimal projection approach to fixed-order dynamic compensation is merged with the guaranteed cost control approach to robust stability and performance to obtain a theory of full- and reduced-order robust control design. The principal result is a sufficient condition for characterizing dynamic controllers of fixed dimension which are guaranteed to provide both robust stability and performance. The sufficient conditions involve a system of modified Riccati and Lyapunov equations coupled by an oblique projection, i.e., idempotent matrix, as well as the uncertainty bounds. Finally, in contrast to the usual separated Riccati equations, the full-order result involves a coupled system consisting of two modified Riccati equations and two modified Lyapunov equations coupled by the uncertainty bounds. The coupling illustrates the breakdown of the separation principle for LQG control with real-valued structured plant parameter variations.

Supported in part by the Air Force Office of Scientific Research under contract F49620-86-C-0002.

1. Introduction

The direct method of Lyapunov has proven to be an effective approach to robust analysis and design of feedback control laws. References [B1], [B2], [BCL], [BG2], [CL], [CP], [ER], [GB], [H], [KB], [KBH], [L], [PH], [TB], [VW] comprise only a representative collection of the increasing literature in this area. In performing robust synthesis there are two principal issues, namely, stability robustness and performance robustness. Stability robustness addresses the problem of guaranteeing stability of the closed-loop system for plant perturbations within a specified class of uncertainties. In addition to guaranteeing robust stability, it is often desirable to minimize the worst-case performance degradation within a given robust stability range. Although both robust stability and performance are of interest in practice, most of the literature involving quadratic Lyapunov functions is confined to the problem of robust stability. A notable exception is the early work of Chang and Peng ([CP]) which also provides bounds on worst-case quadratic performance within full-state feedback control design.

The contribution of the present paper is a methodology for designing controllers which provide both robust stability and robust performance over a prescribed range of real-valued structured plant parameter variations. The feedback law is in the form of a fixed-order (i.e., full- or reduced-order) strictly proper dynamic compensator. The overall approach is based upon the merging of two distinct control-design techniques, namely, the guaranteed cost control approach to robust performance ([CP]) and the optimal projection approach to fixed-order dynamic compensation ([BH1],[HB]). The principal motivation for our approach is to permit greater flexibility in the design of robust feedback laws by providing an alternative to full-state feedback and full-order dynamic compensation.

The guaranteed cost control approach ([CP]) adopted in the present paper utilizes a performance bound to provide robust performance in addition to robust stability. Here, robust performance refers to a guaranteed bound on the worst-case value of the expectation of a quadratic cost criterion over a prescribed uncertainty set. This quadratic criterion is precisely the standard cost functional of linear-quadratic-Gaussian control theory. By bounding the worst-case value of this criterion over a specified range of plant uncertainties, we effectively bound the variances of specified states and control signals.

To bound the worst-case closed-loop performance, we require a bound on the effect of plant uncertainties on the steady-state closed-loop covariance matrix. The form of the guaranteed cost

control bound utilized herein was originally motivated by the effect of multiplicative white noise on the state covariance ([B2],[BG2]). Since this bound is differentiable with respect to the covariance matrix and compensator gains, it permits optimal design via first-order necessary conditions. This approach is not possible using the nondifferentiable bound originally proposed in [CP]. An alternative differentiable bound proposed in [PH] for full-state feedback has been extended to fixed-order dynamic compensation in [BH2].

In the present paper, the guaranteed cost technique is used to bound the closed-loop performance and characterize robustly stabilizing controllers. This performance bound is then interpreted as an *auxiliary cost* which is to be minimized by the choice of compensator gains. The *actual* performance for a given realization of the parameter uncertainty is thus guaranteed to lie below this bound. In the presence of a stabilizability (disturbability) assumption, the robust performance bound automatically implies robust stability. The auxiliary cost and the Lyapunov equation constraint together form the Auxiliary Minimization Problem. Since the Auxiliary Minimization Problem is a nonconvex mathematical programming problem with differentiable data, it is amenable to first-order necessary conditions.

One feature of this approach is that since the necessary conditions are obtained for the Auxiliary Minimization Problem rather than the original problem, extremals are guaranteed to provide both robust stability and performance. Note that this is true for every extremal of the Auxiliary Minimization Problem whether it corresponds to a local minimum, local maximum, or otherwise. Of course, the global minimum is most likely to provide the best worst-case performance over the robust stability range. In any case, necessary conditions for the Auxiliary Minimization Problem effectively serve as *sufficient* conditions for robust stability with a guaranteed performance bound.

The present paper encompasses a rigorous development of sufficient conditions for robust stability and performance via fixed-order dynamic compensation. These sufficient conditions are in the form of a coupled system of algebraic matrix equations consisting of two modified Riccati equations and two modified Lyapunov equations. The coupling is due to the optimal projection, which characterizes reduced-order controllers, and the uncertainty bounds, which account for the effect of parameter uncertainties on the performance functional. When the compensator order is constrained to be equal to the dimension of the plant and the uncertainty bounds are absent, the equations specialize to the usual pair of separated Riccati equations of steady-state LQG theory.

We are quick to point out, however, that our approach is constructive in nature rather than

existential. That is, our sufficient conditions provide explicit formulae for robust, fixed-order feedback gains when the Auxiliary Minimization Problem has a solution. In this sense our constructive conditions can be viewed as complementary to existential results on robust stabilizability. Specifically, the existence of a solution to the Auxiliary Minimization Problem and associated design equations depends upon stabilizability via fixed-order controllers as well as the sharpness of the quadratic Lyapunov bounds. The stabilizability problem has been studied using independent methods (see, e.g., [BHK]), while the conservatism of the bounds is considered in [BH3]. In addition, we state a local existence result for solvability of the design equations which assumes only nominal stabilizability.

The contents and scope of the paper are as follows. In Section 2 we state the robust stability and performance problems for fixed-order dynamic compensation with plant parameter uncertainty. In Section 3 a modified Lyapunov equation is introduced whose solution, when it exists, is guaranteed to bound the steady-state closed-loop covariance over the specified range of plant uncertainty. A performance bound is then given in terms of the covariance bound. In Section 4 we view the performance bound as an auxiliary cost and consider the problem of minimizing the auxiliary cost subject to the modified Lyapunov equation and a definiteness condition as side constraints. These side constraints have the property that all admissible elements provide robust stability and performance (Proposition 4.1). In Section 5 the uncertainty set and bound for constructing the modified Lyapunov equation are given concrete forms. Specifically, the uncertainty set has the form of an ellipsoidal region in parameter space while the modified Lyapunov equation includes additional linear terms to bound the uncertainty. A sufficient condition involving Kronecker sums and products implies the existence of a unique, nonnegative-definite solution to the modified Lyapunov equation. Section 6 presents the first-order necessary conditions (Theorem 6.1) for the Auxiliary Minimization Problem under minor additional technical conditions to ensure the applicability of the Lagrange multiplier technique. As discussed above, these necessary conditions are in the form of extended optimal projection equations. A partial converse of the necessary conditions shows that solutions of these algebraic equations provide, by construction, a solution of the original modified Lyapunov equation. This result is combined in Section 7 (Theorem 7.1) with a stabilizability assumption to guarantee robust stability with a robust performance bound. In addition, we state an existence result for local solvability of the design equations by applying a result from [R1],[R2] (Theorem 7.2). To draw connections with standard LQG theory, in Section 8 we specialize Theorem 7.1 to the full-order case. In contrast to the pair of separated Riccati equations of standard LQG theory,

the full-order result in the presence of plant parameter variations is given by a coupled system of four modified Riccati/Lyapunov equations. In Section 9 the theory is illustrated by means of an example due to Doyle ([D]). This problem was also considered in [BG1] before the robustness theory developed herein was available. Hence the present paper can be viewed as the rigorous mathematical foundation which legitimizes the heretofore ad hoc robustness approach of [BG1].

Notation. Note: All matrices have real entries

| | |
|---|--|
| $\mathbb{R}, \mathbb{R}^{r \times s}, \mathbb{R}^r, \mathbb{E}$ | real numbers, $r \times s$ real matrices, $\mathbb{R}^{r \times 1}$, expected value |
| $I_r, (\)^T, 0_{r \times s}, 0_r$ | $r \times r$ identity matrix, transpose, $r \times s$ zero matrix, $0_{r \times r}$ |
| $Z_{(i,j)}, \text{tr } Z$ | (i,j) -element of matrix Z , trace of square matrix Z |
| \oplus, \otimes | Kronecker sum, Kronecker product ([B3]) |
| $\mathbb{S}^r, \mathbb{I}^r, \mathbb{P}^r$ | $r \times r$ symmetric, nonnegative-definite, positive-definite matrices |
| $Z_1 \leq Z_2, Z_1 < Z_2$ | $Z_2 - Z_1 \in \mathbb{I}^r, Z_2 - Z_1 \in \mathbb{P}^r, Z_1, Z_2 \in \mathbb{S}^r$ |
| $n, m, \ell, n_c, p; \tilde{n}$ | positive integers; $n + n_c$ |
| x, u, y, x_c, \tilde{x} | $n, m, \ell, n_c, \tilde{n}$ -dimensional vectors |
| $A, \Delta A; B, \Delta B$ | $n \times n$ matrices; $n \times m$ matrices |
| $C, \Delta C; D, \Delta D$ | $\ell \times n$ matrices; $\ell \times m$ matrices |
| A_c, B_c, C_c | $n_c \times n_c, n_c \times \ell, m \times n_c$ matrices |
| α | positive number |
| $A_\alpha, A_{c\alpha}$ | $A + \frac{\alpha}{2} I_n, A_c + \frac{\alpha}{2} I_{n_c}$ |
| α_i | positive number, $i = 1, \dots, p$ |
| γ_i | $\alpha_i^2 / \alpha, i = 1, \dots, p$ |
| σ_i | real number, $i = 1, \dots, p$ |
| R_1, R_2 | state, control weighting matrices; $R_1 \in \mathbb{I}^n, R_2 \in \mathbb{P}^m$ |
| R_{12} | $n \times m$ cross-weighting matrix; $R_1 - R_{12} R_2^{-1} R_{12}^T \geq 0$ |
| \tilde{R} | $\begin{bmatrix} R_1 & R_{12} C_c \\ C_c^T R_{12}^T & C_c^T R_2 C_c \end{bmatrix}$ |
| $w_1(\cdot), w_2(\cdot)$ | n, ℓ -dimensional white noise |
| V_1, V_2 | intensity of $w_1(\cdot), w_2(\cdot)$; $V_1 \in \mathbb{I}^n, V_2 \in \mathbb{P}^\ell$ |
| V_{12} | $n \times \ell$ cross intensity of $w_1(\cdot), w_2(\cdot)$ |
| $\tilde{w}(\cdot), \tilde{V}$ | $\begin{bmatrix} w_1(\cdot) \\ B_c w_2(\cdot) \end{bmatrix}, \begin{bmatrix} V_1 & V_{12} B_c^T \\ B_c V_{12}^T & B_c V_2 B_c^T \end{bmatrix}$ |

2. Robust Stability and Robust Performance Problems

In this section we state the Robust Stability Problem and Robust Performance Problem. Both problems involve a set $\mathcal{U} \subset \mathbb{R}^{n \times n} \times \mathbb{R}^{n \times m} \times \mathbb{R}^{l \times n} \times \mathbb{R}^{l \times m}$ of uncertain perturbations $(\Delta A, \Delta B, \Delta C, \Delta D)$ of the nominal system matrices (A, B, C, D) . The goal of the Robust Stability Problem is to determine a fixed-order, strictly proper dynamic compensator (A_c, B_c, C_c) which stabilizes the plant for all variations in \mathcal{U} . In this section and the following section no explicit assumptions are required for the set \mathcal{U} . In Section 5 the structure of variations in \mathcal{U} will be specified.

Robust Stability Problem. For fixed $n_c \leq n$ determine (A_c, B_c, C_c) such that the closed-loop system consisting of the n th-order controlled plant

$$\dot{x}(t) = (A + \Delta A)x(t) + (B + \Delta B)u(t), \quad t \in [0, \infty), \quad (2.1)$$

measurements

$$y(t) = (C + \Delta C)x(t) + (D + \Delta D)u(t), \quad (2.2)$$

and n_c th-order dynamic compensator

$$\dot{x}_c(t) = A_c x_c(t) + B_c y(t), \quad (2.3)$$

$$u(t) = C_c x_c(t), \quad (2.4)$$

is asymptotically stable for all $(\Delta A, \Delta B, \Delta C, \Delta D) \in \mathcal{U}$.

The Robust Performance Problem involves, in addition, white plant disturbances and measurement noise. The goal of this problem is to determine a fixed-order, strictly proper compensator (A_c, B_c, C_c) which minimizes the worst-case value over the uncertainty set \mathcal{U} of a steady-state average quadratic performance criterion.

Robust Performance Problem. For fixed $n_c \leq n$, determine (A_c, B_c, C_c) such that, for the closed-loop system consisting of the n th-order controlled and disturbed plant

$$\dot{x}(t) = (A + \Delta A)x(t) + (B + \Delta B)u(t) + w_1(t), \quad t \in [0, \infty), \quad (2.5)$$

noisy measurements

$$y(t) = (C + \Delta C)x(t) + (D + \Delta D)u(t) + w_2(t), \quad (2.6)$$

and n_c th-order dynamic compensator (2.3), (2.4), the performance criterion

$$J(A_c, B_c, C_c) \triangleq \sup_{(\Delta A, \Delta B, \Delta C, \Delta D) \in \mathcal{U}} \limsup_{t \rightarrow \infty} \mathbb{E}[x^T(t)R_1 x(t) + 2x^T(t)R_{12}u(t) + u^T(t)R_2 u(t)] \quad (2.7)$$

is minimized.

Remark 2.1. The cost functional (2.7) is identical to the standard LQG criterion with the exception of the supremum for evaluating worst-case quadratic performance over \mathcal{U} . Note that (2.7) can also be written in terms of an averaged integral, i.e.,

$$J(A_e, B_e, C_e) = \sup_{(\Delta A, \Delta B, \Delta C, \Delta D) \in \mathcal{U}} \limsup_{t \rightarrow \infty} \frac{1}{t} \mathbb{E} \left\{ \int_0^t [x^T(s) R_1 x(s) + 2x^T(s) R_{12} u(s) + u^T(s) R_2 u(s)] ds \right\}.$$

For practical application, the cost (2.7) provides the means for minimizing the variances of selected state variables and control signals. This can be achieved by appropriate selection of the matrices R_1 and R_2 which serve as design weights. For robust performance the goal is to minimize the worst-case variances of selected variables over the plant uncertainty.

For each uncertain variation $(\Delta A, \Delta B, \Delta C, \Delta D) \in \mathcal{U}$, the undisturbed closed-loop system (2.1)–(2.4) can be written as

$$\dot{\tilde{x}}(t) = (\tilde{A} + \Delta \tilde{A}) \tilde{x}(t), \quad t \in [0, \infty), \quad (2.8)$$

where

$$\tilde{x}(t) \triangleq \begin{bmatrix} x(t) \\ x_c(t) \end{bmatrix}, \quad \tilde{A} \triangleq \begin{bmatrix} A & BC_e \\ B_e C & A_e + B_e DC_e \end{bmatrix}, \quad \Delta \tilde{A} \triangleq \begin{bmatrix} \Delta A & \Delta BC_e \\ B_e \Delta C & B_e \Delta DC_e \end{bmatrix}.$$

Similarly, the disturbed closed-loop system (2.3)–(2.6) can be written as

$$\dot{\tilde{x}}(t) = (\tilde{A} + \Delta \tilde{A}) \tilde{x}(t) + \tilde{w}(t), \quad t \in [0, \infty), \quad (2.9)$$

where the closed-loop disturbance $\tilde{w}(t)$ has intensity $\tilde{V} \in \mathbb{N}^{\tilde{n}}$.

3. Sufficient Conditions for Robust Stability and Performance

In practice, steady-state performance is only of interest when the undisturbed closed-loop system (2.8) is robustly stable over \mathcal{U} . The following result, which expresses the performance in terms of the steady-state closed-loop second-moment matrix, is immediate.

Lemma 3.1. Let (A_e, B_e, C_e) be given and suppose the system (2.8) is stable for all $(\Delta A, \Delta B, \Delta C, \Delta D) \in \mathcal{U}$. Then

$$J(A_e, B_e, C_e) = \sup_{(\Delta A, \Delta B, \Delta C, \Delta D) \in \mathcal{U}} \text{tr } \tilde{Q}_{\Delta \tilde{A}} \tilde{R}, \quad (3.1)$$

where $\tilde{Q}_{\Delta\tilde{A}} \triangleq \lim_{t \rightarrow \infty} \mathbb{E}[\tilde{x}(t)\tilde{x}^T(t)] \in \mathbb{N}^{\tilde{n}}$ is the unique solution to

$$0 = (\tilde{A} + \Delta\tilde{A})\tilde{Q}_{\Delta\tilde{A}} + \tilde{Q}_{\Delta\tilde{A}}(\tilde{A} + \Delta\tilde{A})^T + \tilde{V}. \quad (3.2)$$

Remark 3.1. When \mathcal{U} is compact, "sup" in (3.1) can be replaced by "max".

The key step in guaranteeing robust stability and performance is to replace the uncertain terms in the covariance Lyapunov equation (3.2) by a bounding function Ω . Note that since $\Delta\tilde{A}$ is independent of A_c , the bounding function need only depend upon B_c and C_c .

Theorem 3.1. Let $\Omega : \mathbb{N}^{\tilde{n}} \times \mathbb{R}^{n_c \times \ell} \times \mathbb{R}^{m \times n_c} \rightarrow \mathbb{S}^{\tilde{n}}$ be such that

$$\begin{aligned} \Delta\tilde{A}\mathcal{Q} + \mathcal{Q}\Delta\tilde{A}^T &\leq \Omega(\mathcal{Q}, B_c, C_c), \\ (\Delta A, \Delta B, \Delta C, \Delta D) &\in \mathcal{U}, \quad (\mathcal{Q}, B_c, C_c) \in \mathbb{N}^{\tilde{n}} \times \mathbb{R}^{n_c \times \ell} \times \mathbb{R}^{m \times n_c}, \end{aligned} \quad (3.3)$$

and, for given (A_c, B_c, C_c) , suppose there exists $\mathcal{Q} \in \mathbb{N}^{\tilde{n}}$ satisfying

$$0 = \tilde{A}\mathcal{Q} + \mathcal{Q}\tilde{A}^T + \Omega(\mathcal{Q}, B_c, C_c) + \tilde{V}. \quad (3.4)$$

Then

$$\begin{aligned} (\tilde{A} + \Delta\tilde{A}, [\tilde{V} + \Omega(\mathcal{Q}, B_c, C_c) - (\Delta\tilde{A}\mathcal{Q} + \mathcal{Q}\Delta\tilde{A}^T)]^{\frac{1}{2}}) &\text{ is stabilizable,} \\ (\Delta A, \Delta B, \Delta C, \Delta D) &\in \mathcal{U}, \end{aligned} \quad (3.5)$$

if and only if

$$\tilde{A} + \Delta\tilde{A} \text{ is asymptotically stable, } (\Delta A, \Delta B, \Delta C, \Delta D) \in \mathcal{U}. \quad (3.6)$$

In this case,

$$\tilde{Q}_{\Delta\tilde{A}} \leq \mathcal{Q}, \quad (\Delta A, \Delta B, \Delta C, \Delta D) \in \mathcal{U}, \quad (3.7)$$

where $\tilde{Q}_{\Delta\tilde{A}}$ is given by (3.2), and

$$J(A_c, B_c, C_c) \leq \text{tr } \mathcal{Q}\tilde{R}. \quad (3.8)$$

Proof. First note for clarity that in (3.3) \mathcal{Q} denotes an arbitrary element of $\mathbb{N}^{\tilde{n}}$ since (3.3) holds for all $\mathcal{Q} \in \mathbb{N}^{\tilde{n}}$, while in (3.4) \mathcal{Q} denotes a specific solution to (3.4). Now for $(\Delta A, \Delta B, \Delta C, \Delta D) \in \mathcal{U}$, (3.4) is equivalent to

$$0 = (\tilde{A} + \Delta\tilde{A})\mathcal{Q} + \mathcal{Q}(\tilde{A} + \Delta\tilde{A})^T + \Omega(\mathcal{Q}, B_c, C_c) - (\Delta\tilde{A}\mathcal{Q} + \mathcal{Q}\Delta\tilde{A}^T) + \tilde{V}. \quad (3.9)$$

Hence, by assumption, (3.9) has a solution $\mathcal{Q} \in \mathbb{N}^{\tilde{n}}$ for all $(\Delta A, \Delta B, \Delta C, \Delta D) \in \mathcal{U}$ and, by (3.3), $\Omega(\mathcal{Q}, B_c, C_c) - (\Delta\tilde{A}\mathcal{Q} + \mathcal{Q}\Delta\tilde{A}^T)$ is nonnegative definite. Now if the stabilizability condition

(3.5) holds for all $(\Delta A, \Delta B, \Delta C, \Delta D) \in \mathcal{U}$, it follows from Lemma 12.2 of [W] that $\tilde{A} + \Delta \tilde{A}$ is asymptotically stable for all $(\Delta A, \Delta B, \Delta C, \Delta D) \in \mathcal{U}$. Conversely, if $\tilde{A} + \Delta \tilde{A}$ is asymptotically stable for all $(\Delta A, \Delta B, \Delta C, \Delta D) \in \mathcal{U}$, then (3.5) holds. Next, subtracting (3.2) from (3.9) yields

$$0 = (\tilde{A} + \Delta \tilde{A})(Q - \tilde{Q}_{\Delta \tilde{A}}) + (Q - \tilde{Q}_{\Delta \tilde{A}})(\tilde{A} + \Delta \tilde{A})^T + \Omega(Q, B_c, C_c) - (\Delta \tilde{A}Q + Q\Delta \tilde{A}^T),$$

or, equivalently, since $\tilde{A} + \Delta \tilde{A}$ is asymptotically stable for all $(\Delta A, \Delta B, \Delta C, \Delta D) \in \mathcal{U}$,

$$Q - \tilde{Q}_{\Delta \tilde{A}} = \int_0^\infty e^{(\tilde{A} + \Delta \tilde{A})t} [\Omega(Q, B_c, C_c) - (\Delta \tilde{A}Q + Q\Delta \tilde{A}^T)] e^{(\tilde{A} + \Delta \tilde{A})^T t} dt \geq 0,$$

which implies (3.7). The performance bound (3.8) is now an immediate consequence of (3.7). \square

Remark 3.2. In applying Theorem 3.1 it may be convenient to replace condition (3.5) with a stronger condition which is easier to verify in practice. Clearly, (3.5) is satisfied if $\tilde{V} + \Omega(Q, B_c, C_c) - (\Delta \tilde{A}Q + Q\Delta \tilde{A}^T)$ is positive definite for all $(\Delta A, \Delta B, \Delta C, \Delta D) \in \mathcal{U}$. This will be the case, for example, if either \tilde{V} is positive definite or strict inequality holds in (3.3). Also, it follows from Theorem 3.6 of [W] that (3.5) is implied by the stronger condition

$$(\tilde{A} + \Delta \tilde{A}, \tilde{V}^{\frac{1}{2}}) \text{ is stabilizable, } (\Delta A, \Delta B, \Delta C, \Delta D) \in \mathcal{U}. \quad (3.10)$$

Remark 3.3. The covariance bound (3.7) can also be used to analyze the effect of disturbances on specified state variables. For example, if $E_1 \in \mathbb{R}^{q \times n}$ then (3.7) implies

$$[E_1 \ 0_{q \times n_c}] \tilde{Q}_{\Delta \tilde{A}} \begin{bmatrix} E_1^T \\ 0_{n_c \times q} \end{bmatrix} \leq [E_1 \ 0_{q \times n_c}] Q \begin{bmatrix} E_1^T \\ 0_{n_c \times q} \end{bmatrix} \quad (3.11)$$

so that the right hand side of (3.11) serves as a bound on selected state variances. For control-design purposes we effectively set $R_1 = E_1^T E_1$. Similar remarks apply to obtaining bounds on the variances of control signals.

4. The Auxiliary Minimization Problem

The key step in our development involves consideration of the performance bound (3.8) in place of the actual worst-case performance $J(A_c, B_c, C_c)$. This leads to the following problem.

Auxiliary Minimization Problem. Determine (Q, A_c, B_c, C_c) which minimizes

$$J(Q, A_c, B_c, C_c) \triangleq \text{tr } Q\tilde{R} \quad (4.1)$$

subject to (3.4) and

$$Q \in \mathbb{N}^{\tilde{n}}. \quad (4.2)$$

The relationship between the Auxiliary Minimization Problem and the Robust Stability and Performance Problems is straightforward as shown by the following observation.

Proposition 4.1. If (Q, A_c, B_c, C_c) satisfies (3.4), (4.2) and the stabilizability condition (3.5) holds, then $\tilde{A} + \Delta\tilde{A}$ is asymptotically stable for all $(\Delta A, \Delta B, \Delta C, \Delta D) \in \mathcal{U}$, and

$$J(A_c, B_c, C_c) \leq J(Q, A_c, B_c, C_c). \quad (4.3)$$

Proof. Since (3.4) has a solution $Q \in \mathbb{R}^{\tilde{n}}$ and the stabilizability condition (3.5) holds, the hypotheses of Theorem 3.1 are satisfied so that robust stability with robust performance bound (3.8) is guaranteed. Note that (4.3) is merely a restatement of (3.8). \square

Several comments are in order. Since the *auxiliary cost* (4.1) is an upper bound for the actual cost (2.7), it is clearly desirable to minimize (4.1) over Q and the controller gains. Note, however, that the Auxiliary Minimization Problem is a nonconvex mathematical programming problem on a noncompact set. Hence guarantees of existence of solutions and sufficient conditions for global optimality cannot be obtained without imposing additional confining assumptions. To develop nonrestrictive results, we proceed in Section 6 by deriving necessary conditions for optimality which require no further assumptions except that Ω be differentiable and that the minimization be performed over an open set. In the next section we construct a bound Ω which possesses the required smoothness.

5. Uncertainty Structure and the Guaranteed Cost Bound

Having established the theoretical basis for our approach, we now assign explicit structure to the set \mathcal{U} and bounding function Ω . Specifically, the uncertainty set \mathcal{U} is assumed to be of the form

$$\begin{aligned} \mathcal{U} = \{(\Delta A, \Delta B, \Delta C, \Delta D) : \Delta A = \sum_{i=1}^p \sigma_i A_i, \Delta B = \sum_{i=1}^p \sigma_i B_i, \Delta C = \sum_{i=1}^p \sigma_i C_i, \\ \Delta D = \sum_{i=1}^p \sigma_i D_i, \sum_{i=1}^p \sigma_i^2 / \alpha_i^2 \leq 1\}, \end{aligned} \quad (5.1)$$

where, for $i = 1, \dots, p$: $A_i \in \mathbb{R}^{n \times n}$, $B_i \in \mathbb{R}^{n \times m}$, $C_i \in \mathbb{R}^{l \times n}$, and $D_i \in \mathbb{R}^{l \times m}$ are fixed matrices denoting the structure of the parametric uncertainty; α_i is a given positive number; and σ_i is an uncertain real parameter. Note that the uncertain parameters σ_i are assumed to lie in a specified ellipsoidal region in \mathbb{R}^p . The closed-loop system (2.8) thus has structured uncertainty of the form

$$\Delta\tilde{A} = \sum_{i=1}^p \sigma_i \tilde{A}_i, \quad (5.2)$$

where

$$\tilde{A}_i \triangleq \begin{bmatrix} A_i & B_i C_c \\ B_c C_i & B_c D_i C_c \end{bmatrix}, \quad i = 1, \dots, p.$$

The uncertainty set \mathcal{U} is general in the sense that no explicit assumptions such as the matching conditions used in [BCL] will be made with regard to the structure of A_i, B_i, C_i , and D_i . We do, however, require (as is evident from (5.1)) that the uncertain parameters σ_i appear linearly in the off-nominal perturbations which is more confining than matching assumptions. Note that the symmetry of the uncertainty set entails no loss of generality by requiring only a redefinition of the nominal plant matrices.

In order to obtain *explicit* gain expressions for (A_c, B_c, C_c) in Section 6, we shall require one additional technical assumption. Specifically, we assume that for each $i \in \{1, \dots, p\}$, at most one of the matrices B_i, C_i , and D_i is nonzero. This condition thus implies that a given uncertain parameter σ_i may appear explicitly in both ΔA and ΔB , or both ΔA and ΔC , or both ΔA and ΔD , or only ΔA , but not (say) in both ΔB and ΔD . Thus we can account partially (but not totally) for *correlated* parameter uncertainties in different plant matrices. If a given uncertain parameter does arise in both (say) ΔB and ΔD , then it must be represented by two distinct uncertain parameters. If this assumption is not imposed, then optimality conditions can still be derived, but at the expense of closed-form gain expressions.

For the structure of \mathcal{U} as specified by (5.1), the bound Ω satisfying (3.3) can now be given a concrete form.

Proposition 5.1. Let α be an arbitrary positive scalar. Then the function

$$\Omega(Q, B_c, C_c) = \alpha Q + \alpha^{-1} \sum_{i=1}^p \alpha_i^2 \tilde{A}_i Q \tilde{A}_i^T \quad (5.3)$$

satisfies (3.3) with \mathcal{U} given by (5.1).

Proof. Note that

$$\begin{aligned} 0 &\leq \sum_{i=1}^p [(\alpha^{\frac{1}{2}} \sigma_i / \alpha_i) I_n - (\alpha_i / \alpha^{\frac{1}{2}}) \tilde{A}_i] Q [(\alpha^{\frac{1}{2}} \sigma_i / \alpha_i) I_n - (\alpha_i / \alpha^{\frac{1}{2}}) \tilde{A}_i]^T \\ &= \alpha \sum_{i=1}^p (\sigma_i^2 / \alpha_i^2) Q + \alpha^{-1} \sum_{i=1}^p \alpha_i^2 \tilde{A}_i Q \tilde{A}_i^T - \sum_{i=1}^p \sigma_i (\tilde{A}_i Q + Q \tilde{A}_i^T), \end{aligned}$$

which, since $\sum_{i=1}^p \sigma_i^2 / \alpha_i^2 \leq 1$, implies (3.3). \square

Remark 5.1. Note that the bound Ω given by (5.3) consists of two distinct terms. The first term αQ can be thought of as arising from an exponential time weighting of the cost, or, equivalently, from a uniform right shift of the open-loop dynamics ([AM]). The second term $\alpha^{-1} \sum_{i=1}^p \alpha_i^2 \tilde{A}_i Q \tilde{A}_i^T$ arises naturally from a multiplicative white noise model ([BG1],[BG2],[B2]). Such interpretations have no bearing on the results obtained here since only the bound Ω defined by (5.3) is required. Note that the bound (5.3) is valid for all positive α . A similar bound was also considered in [KB].

With Ω defined by (5.3), the modified Lyapunov equation (3.4) becomes

$$0 = \tilde{A} Q + Q \tilde{A}^T + \alpha Q + \alpha^{-1} \sum_{i=1}^p \alpha_i^2 \tilde{A}_i Q \tilde{A}_i^T + \tilde{V} \quad (5.4)$$

or, equivalently,

$$0 = \tilde{A}_\alpha Q + Q \tilde{A}_\alpha^T + \sum_{i=1}^p \gamma_i \tilde{A}_i Q \tilde{A}_i^T + \tilde{V}, \quad (5.5)$$

where

$$\tilde{A}_\alpha \triangleq \tilde{A} + \frac{\alpha}{2} I_{\tilde{n}} = \begin{bmatrix} A_\alpha & B C_c \\ B_c C & A_{c\alpha} + B_c D C_c \end{bmatrix} \quad (5.6)$$

and $\gamma_i \triangleq \alpha_i^2 / \alpha$. Note that (5.5) is equivalent to

$$0 = A \text{vec } Q + \text{vec } \tilde{V}, \quad (5.7)$$

where "vec" is the column-stacking operation defined in [B3] and A is defined by

$$A \triangleq \tilde{A}_\alpha \oplus \tilde{A}_\alpha + \sum_{i=1}^p \gamma_i \tilde{A}_i \otimes \tilde{A}_i.$$

Next, using the bound Ω given by (5.3) and \mathcal{U} given by (5.1) we present a result which guarantees the existence of a nonnegative-definite solution to (3.4) or, equivalently, (5.5) for a given controller (A_c, B_c, C_c) . For the converse we view \tilde{V} as an arbitrary element of $\mathbb{R}^{\tilde{n}}$.

Proposition 5.2. Let (A_c, B_c, C_c) be given and let $\alpha > 0$. If A is asymptotically stable, then there exists a unique $\tilde{n} \times \tilde{n}$ Q satisfying (5.5) and, furthermore, $Q \geq 0$. Conversely, if for all $\tilde{V} \in \mathbb{R}^{\tilde{n}}$ there exists $Q \geq 0$ satisfying (5.5), then A is asymptotically stable.

Proof. Since (5.5) is equivalent to

$$Q = -\text{vec}^{-1}[A^{-1} \text{vec } \tilde{V}], \quad (5.8)$$

existence and uniqueness hold. To prove that Q is nonnegative definite, we rewrite (5.8) as

$$Q = \int_0^\infty \text{vec}^{-1}[e^{At} \text{vec } \tilde{V}] dt \quad (5.9)$$

and show that the integrand is nonnegative definite for all $t \in [0, \infty)$. [Note that the following argument does not require that A be stable.] Using the Lie exponential product formula, the exponential in (5.9) can be written as

$$e^{At} = \lim_{k \rightarrow \infty} \left\{ \exp\left[\frac{1}{k}(\tilde{A}_\alpha \oplus \tilde{A}_\alpha)t\right] \prod_{i=1}^p \exp\left[\frac{1}{k}\gamma_i(\tilde{A}_i \otimes \tilde{A}_i)t\right] \right\}^k. \quad (5.10)$$

For convenience, let S and N be $r \times r$ matrices with $N \geq 0$. Since (see [B3])

$$\text{vec}^{-1}[(S \otimes S) \text{vec } N] = S N S^T \geq 0 \quad (5.11)$$

and

$$(S^k \otimes S^k)(S \otimes S) = S^{k+1} \otimes S^{k+1}, \quad (5.12)$$

it follows that

$$\text{vec}^{-1}[e^{S \otimes S} \text{vec } N] = \sum_{k=0}^{\infty} (k!)^{-1} S^k N S^{kT} \geq 0. \quad (5.13)$$

Furthermore,

$$\text{vec}^{-1}[e^{S \otimes S} \text{vec } N] = \text{vec}^{-1}[(e^S \otimes e^S) \text{vec } N] = e^S N e^{S^T} \geq 0. \quad (5.14)$$

Applying (5.13) and (5.14) alternately with (5.10) and using induction on k it follows that the integrand of (5.9) is nonnegative definite. To prove the converse, note that it follows from (5.5) that Q satisfies

$$Q = \text{vec}^{-1}[e^{At} \text{vec } Q] + \int_0^t \text{vec}^{-1}[e^{As} \text{vec } \tilde{V}] ds, \quad t \in [0, \infty). \quad (5.15)$$

Since the integral term on the right hand side of (5.15) is nonnegative definite, is bounded from above by Q , and $\tilde{V} \in \mathbb{R}^{\tilde{n}}$ is arbitrary, it follows that A is asymptotically stable. \square

Proposition 5.2 shows that a solution to (5.5) exists so long as $\alpha_1, \dots, \alpha_p$ are sufficiently small that A remains stable for some $\alpha > 0$. The following result characterizes values of $\alpha_1, \dots, \alpha_p$ for which A is asymptotically stable. Let $\|\cdot\|$ denote an arbitrary vector norm and its induced matrix norm.

Proposition 5.3. Let (A_c, B_c, C_c) be given, assume \tilde{A} is asymptotically stable, and let $\alpha, \alpha_1, \dots, \alpha_p > 0$. If

$$\|(\tilde{A} \oplus \tilde{A})^{-1}(\alpha I_{\tilde{n}} + \sum_{i=1}^p \gamma_i \tilde{A}_i \otimes \tilde{A}_i)\| < 1, \quad (5.16)$$

then there exists $Q \in \mathbb{R}^n$ satisfying (5.5) and A is asymptotically stable.

Proof. Define $\{Q_k\}_{k=0}^\infty$ where Q_0 satisfies

$$0 = \tilde{A}Q_0 + Q_0\tilde{A}^T + \tilde{V},$$

and Q_{k+1} satisfies

$$0 = \tilde{A}Q_{k+1} + Q_{k+1}\tilde{A}^T + \Omega(Q_k, B_c, C_c) + \tilde{V}.$$

Note that $Q_k \geq 0$, $k = 1, 2, \dots$. Hence it follows that

$$\text{vec } Q_{k+1} - \text{vec } Q_k = -(\tilde{A} \oplus \tilde{A})^{-1} [\text{vec } \Omega(Q_k, B_c, C_c) - \text{vec } \Omega(Q_{k-1}, B_c, C_c)]$$

and thus

$$\|\text{vec } Q_{k+1} - \text{vec } Q_k\| \leq \|(\tilde{A} \oplus \tilde{A})^{-1}(\alpha I_{\tilde{n}^2} + \sum_{i=1}^p \gamma_i \tilde{A}_i \otimes \tilde{A}_i)\| \|\text{vec } Q_k - \text{vec } Q_{k-1}\|.$$

Using (5.16) it follows that $Q \triangleq \lim_{k \rightarrow \infty} Q_k$ exists. Thus $Q \geq 0$ and satisfies (5.5). Furthermore, since $\tilde{V} \in \mathbb{R}^n$ can be considered arbitrary, Proposition 5.2 implies that A is asymptotically stable. \square

6. Necessary Conditions for the Auxiliary Minimization Problem

The derivation of the necessary conditions for the Auxiliary Minimization Problem is based upon the Fritz John form of the Lagrange multiplier theorem. Rigorous application of this theorem requires that we further restrict (Q, A_c, B_c, C_c) to the open set

$$S \triangleq \{(Q, A_c, B_c, C_c) : Q \in \mathbb{P}^n, A \text{ is asymptotically stable,} \\ \text{and } (A_c, B_c, C_c) \text{ is controllable and observable}\}.$$

As will be seen, the constraint $(Q, A_c, B_c, C_c) \in S$ is not required for either robust stability or robust performance since Proposition 4.1 shows that only (3.4), (3.5) and (4.2) are needed. Rather, the set S constitutes sufficient conditions under which the Lagrange multiplier technique is applicable to the Auxiliary Minimization Problem. Specifically, the condition $Q > 0$ replaces (4.2) by an open set constraint, the asymptotic stability of A serves as a normality condition which further implies that the dual P of Q is nonnegative definite, and (A_c, B_c, C_c) minimal is a nondegeneracy condition which implies that the lower right $n_c \times n_c$ subblocks of Q and P are positive definite thus yielding explicit expressions for B_c and C_c . Note that by Proposition 5.2 the condition that A be

asymptotically stable also implies that (5.5) has a unique, nonnegative solution. Finally, we point out that the stabilizability condition (3.5) and stability condition (3.6) play no role in determining solutions of the Auxiliary Minimization Problem.

In order to state the main results we require some additional notation and a lemma concerning pairs of nonnegative-definite matrices. For a real $n \times n$ matrix Z define the set of real diagonalizing matrices

$$\mathcal{D}(Z) \triangleq \{\Psi \in \mathbb{R}^{n \times n} : \Psi^{-1}Z\Psi \text{ is diagonal}\},$$

and, for a pair of $n \times n$ symmetric matrices X, Y define the set of real *contragrediently diagonalizing* matrices

$$\mathcal{C}(X, Y) \triangleq \{\Psi \in \mathcal{D}(XY) : \Psi^{-1}X\Psi^{-T} \text{ and } \Psi^TY\Psi \text{ are diagonal}\}$$

and the subset of real *balancing* transformations

$$\mathcal{B}(X, Y) \triangleq \{\Psi \in \mathcal{C}(X, Y) : \Psi^{-1}X\Psi^{-T} = \Psi^TY\Psi\}.$$

Of course, a necessary condition for $\mathcal{B}(X, Y)$ to be nonempty is that X, Y , and XY all have the same rank. Note that in general

$$\mathcal{B}(X, Y) \subset \mathcal{C}(X, Y) \subset \mathcal{D}(XY). \quad (6.1)$$

Obviously, a diagonalizable matrix is either invertible (has no zero eigenvalues) or has semisimple zero eigenvalues. Hence if $\mathcal{D}(Z) \neq \emptyset$ then the group generalized inverse $Z^\#$ exists as a special case of the Drazin generalized inverse ([CM]). Note that we limit our consideration to diagonalizable matrices with real eigenvalues. Also, note that there is no assumption here that Z is symmetric. Of course, when Z is symmetric the group, Drazin, and Moore-Penrose generalized inverses coincide.

Lemma 6.1. Let $\hat{Q}, \hat{P} \in \mathbb{R}^n$ and let $r = \text{rank } \hat{Q}\hat{P}$. Then the following statements hold:

- (i) $\hat{Q}\hat{P}$ has nonnegative eigenvalues.
- (ii) $\mathcal{C}(\hat{Q}, \hat{P}) \neq \emptyset$.
- (iii) $\hat{Q}\hat{P}$ is diagonalizable.
- (iv) The $n \times n$ matrix

$$r \triangleq \hat{Q}\hat{P}(\hat{Q}\hat{P})^\# = (\hat{Q}\hat{P})^\#\hat{Q}\hat{P} \quad (6.2)$$

is idempotent, i.e., τ is an oblique projection, and

$$\text{rank } \tau = r. \quad (6.3)$$

(v) There exists $G, \Gamma \in \mathbb{R}^{r \times n}$ and invertible $M \in \mathbb{R}^{r \times r}$ such that

$$\hat{Q}\hat{P} = G^T M \Gamma, \quad (6.4)$$

$$\Gamma G^T = I_r. \quad (6.5)$$

(vi) If $G, \Gamma \in \mathbb{R}^{r \times n}$ and $M \in \mathbb{R}^{r \times r}$ satisfy (6.4) and (6.5), then

$$\text{rank } G = \text{rank } \Gamma = \text{rank } M = r, \quad (6.6)$$

$$(\hat{Q}\hat{P})^\# = G^T M^{-1} \Gamma, \quad (6.7)$$

$$\tau = G^T \Gamma, \quad (6.8)$$

$$\tau G^T = G^T, \quad \Gamma \tau = \Gamma. \quad (6.9)$$

(vii) The matrices $G, \Gamma \in \mathbb{R}^{r \times n}$ and $M \in \mathbb{R}^{r \times r}$ satisfying (6.4) and (6.5) are unique except for a change of basis in \mathbb{R}^r . Furthermore, all such M are diagonalizable with positive eigenvalues.

(viii) If $\text{rank } \hat{Q} = \text{rank } \hat{P} = r$ then $\mathcal{B}(\hat{Q}, \hat{P}) \neq \emptyset$ and

$$\hat{Q} = \tau \hat{Q} = \hat{Q} \tau^T = \tau \hat{Q} \tau^T, \quad (6.10)$$

$$\hat{P} = \tau^T \hat{P} = \hat{P} \tau = \tau^T \hat{P} \tau. \quad (6.11)$$

Proof. See Appendix A. \square

A triple (G, M, Γ) satisfying (6.4) and (6.5) with $G, \Gamma \in \mathbb{R}^{r \times n}$, $M \in \mathbb{R}^{r \times r}$, and $r = \text{rank } \hat{Q}\hat{P}$ will be called a *projective factorization* of $\hat{Q}\hat{P}$. In particular, we shall set $r = n_c$. Furthermore, define the complementary projection

$$\tau_\perp \triangleq I_n - \tau, \quad (6.12)$$

and, for arbitrary $Q, P, \hat{Q}, \hat{P} \in \mathbb{R}^{n \times n}$, $G, \Gamma \in \mathbb{R}^{n \times n}$, $B_c \in \mathbb{R}^{n \times \ell}$, $C_c \in \mathbb{R}^{m \times n}$, and $\alpha > 0$, define the following notation:

$$\begin{aligned} V_{2s} &\triangleq V_2 + \sum_{i=1}^p \gamma_i [C_i(Q + \hat{Q})C_i^T + D_i C_c \Gamma \hat{Q} \Gamma^T C_c^T D_i^T], \\ R_{2s} &\triangleq R_2 + \sum_{i=1}^p \gamma_i [B_i^T(P + \hat{P})B_i + D_i^T B_c^T G \hat{P} G^T B_c D_i], \\ Q_s &\triangleq Q C^T + V_{12} + \sum_{i=1}^p \gamma_i [A_i(Q + \hat{Q})C_i^T + A_i \hat{Q} \Gamma^T C_c^T D_i^T], \\ P_s &\triangleq B^T P + R_{12}^T + \sum_{i=1}^p \gamma_i [B_i^T(P + \hat{P})A_i - D_i^T B_c^T G \hat{P} A_i], \\ A_Q &\triangleq A_\alpha - Q_s V_{2s}^{-1} C, \quad A_P \triangleq A_\alpha - B R_{2s}^{-1} P_s. \end{aligned}$$

The above definitions are for convenience in stating the necessary conditions for the Auxiliary Minimization Problem. This result provides explicit formulae for extremals (Q, A_c, B_c, C_c) of the Auxiliary Minimization Problem. A partial converse shows that this form of the necessary conditions represents no loss of generality with regard to the constraint equation (5.5).

Theorem 6.1. (I) Suppose $(Q, A_c, B_c, C_c) \in S$ solves the Auxiliary Minimization Problem with \mathcal{U} given by (5.1) and Ω given by (5.3). Then there exist $Q, P, \hat{Q}, \hat{P} \in \mathbb{R}^n$ such that, for some projective factorization (G, M, Γ) of $\hat{Q}\hat{P}$, (Q, A_c, B_c, C_c) are given by

$$Q = \begin{bmatrix} Q + \hat{Q} & \hat{Q} \Gamma^T \\ \Gamma \hat{Q} & \Gamma \hat{Q} \Gamma^T \end{bmatrix}, \quad (6.13)$$

$$A_c = \Gamma(A - B R_{2s}^{-1} P_s - Q_s V_{2s}^{-1} C + Q_s V_{2s}^{-1} D R_{2s}^{-1} P_s) G^T, \quad (6.14)$$

$$B_c = \Gamma Q_s V_{2s}^{-1}, \quad (6.15)$$

$$C_c = -R_{2s}^{-1} P_s G^T, \quad (6.16)$$

and such that Q, P, \hat{Q} , and \hat{P} satisfy

$$\begin{aligned} 0 = A_\alpha Q + Q A_\alpha^T + V_1 + \sum_{i=1}^p \gamma_i [A_i Q A_i^T + (A_i - B_i R_{2s}^{-1} P_s) \hat{Q} (A_i - B_i R_{2s}^{-1} P_s)^T \\ - Q_s V_{2s}^{-1} Q_s^T + \tau_\perp Q_s V_{2s}^{-1} Q_s^T \tau_\perp^T], \end{aligned} \quad (6.17)$$

$$\begin{aligned} 0 = A_\alpha^T P + P A_\alpha + R_1 + \sum_{i=1}^p \gamma_i [A_i^T P A_i + (A_i - Q_s V_{2s}^{-1} C_i)^T \hat{P} (A_i - Q_s V_{2s}^{-1} C_i)] \\ - P_s^T R_{2s}^{-1} P_s + \tau_\perp^T P_s^T R_{2s}^{-1} P_s \tau_\perp, \end{aligned} \quad (6.18)$$

$$0 = A_P \hat{Q} + \hat{Q} A_P^T + Q_s V_{2s}^{-1} Q_s^T - \tau_\perp Q_s V_{2s}^{-1} Q_s^T \tau_\perp^T, \quad (6.19)$$

$$0 = A_Q^T \hat{P} + \hat{P} A_Q + P_s^T R_{2s}^{-1} P_s - \tau_\perp^T P_s^T R_{2s}^{-1} P_s \tau_\perp, \quad (6.20)$$

$$\text{rank } \hat{Q} = \text{rank } \hat{P} = \text{rank } \hat{Q}\hat{P} = n_c. \quad (6.21)$$

Furthermore, the auxiliary cost is given by

$$J(Q, A_c, B_c, C_c) = \text{tr}[(Q + \hat{Q})R_1 - 2R_{12}R_{2c}^{-1}P_c\hat{Q} + P_c^TR_{2c}^{-1}R_2R_{2c}^{-1}P_c\hat{Q}]. \quad (6.22)$$

(II) Conversely, if there exist $Q, P, \hat{Q}, \hat{P} \in \mathbb{N}^n$ satisfying (6.17)–(6.21) with B_c and C_c given by (6.15) and (6.16), then (Q, A_c, B_c, C_c) given by (6.13)–(6.16) satisfy (4.2) and (5.5) with $J(Q, A_c, B_c, C_c)$ given by (6.22).

Proof. See Appendix B. \square

Remark 6.1. Theorem 6.1 presents necessary conditions for the Auxiliary Minimization Problem which explicitly characterize extremal quadruples (Q, A_c, B_c, C_c) . These necessary conditions consist of a system of two modified Riccati equations and two modified Lyapunov equations coupled by both the optimal projection τ and uncertainty bounds. If the uncertainty bounds are deleted, then the results of [HB] are recovered.

Remark 6.2. When solving (6.17)–(6.21) numerically, the uncertainty terms can be adjusted to examine tradeoffs between performance and robustness. Specifically, the bounds α_i and the structure matrices A_i, B_i, C_i , and D_i appearing in V_{2s}, R_{2s}, Q_s , and P_s can be varied systematically to determine the region of solvability of (6.17)–(6.21).

Remark 6.3. Although equations (6.17)–(6.21) appear formidable, they are, in fact, quite numerically tractable. For related problems involving coupled Riccati equations, homotopic continuation methods have been shown to be effective ([KLJ],[MB]). Similar algorithms for solving (6.17)–(6.21) have been developed in [GH],[R], while iterative algorithms are discussed in [G2],[GV],[CY].

Remark 6.4. Because of the presence of B_c and C_c in the definitions of V_{2s}, R_{2s}, Q_s , and P_s , the optimality conditions (6.17)–(6.20) are coupled with the gain expressions (6.15) and (6.16) for B_c and C_c . When the problem is specialized to the case $D_i = 0, i = 1, \dots, p$, this coupling disappears and equations (6.17)–(6.20) can be solved without reference to the gain expressions (6.15) and (6.16).

7. Sufficient Conditions for Robust Stability and Performance

In this section we combine Theorem 3.1 with Theorem 6.1 (II) to obtain our main result guaranteeing robust stability and performance.

Theorem 7.1. Suppose there exist $Q, P, \hat{Q}, \hat{P} \in \mathbb{R}^n$ satisfying (6.17)–(6.21) with B_c and C_c given by (6.15) and (6.16). Then, with (Q, A_c, B_c, C_c) given by (6.13)–(6.16), $(\tilde{A} + \Delta\tilde{A}, [\tilde{V} + \Omega(Q, B_c, C_c) - (\Delta\tilde{A}Q + Q\Delta\tilde{A}^T)]^{\frac{1}{2}})$ is stabilizable for all $(\Delta A, \Delta B, \Delta C, \Delta D) \in \mathcal{U}$ if and only if $\tilde{A} + \Delta\tilde{A}$ is asymptotically stable for all $(\Delta A, \Delta B, \Delta C, \Delta D) \in \mathcal{U}$. In this case, the performance (2.7) of the closed-loop system (2.9) satisfies the bound

$$J(A_c, B_c, C_c) \leq \text{tr}[(Q + \hat{Q})R_1 - 2R_{12}R_{2c}^{-1}P_c\hat{Q} + P_c^TR_{2c}^{-1}R_2R_{2c}^{-1}P_c\hat{Q}]. \quad (7.1)$$

Proof. The converse portion of Theorem 6.1 implies that Q given by (6.13) is nonnegative definite and satisfies (5.5) or, equivalently, (3.4). It now follows from Theorem 3.1 that the stabilizability condition (3.5) is equivalent to the asymptotic stability of $\tilde{A} + \Delta\tilde{A}$ for all $(\Delta A, \Delta B, \Delta C, \Delta D) \in \mathcal{U}$. In this case Proposition 4.1 yields robust stability and performance. The robust performance bound (7.1) is a restatement of (4.3) utilizing (6.22). \square

Note that Theorem 7.1 is constructive in nature rather than existential. Specifically, Theorem 7.1 involves a coupled system of modified Riccati/Lyapunov equations (6.17)–(6.21) whose solutions, when they exist, are used to explicitly construct the dynamic feedback gains (6.14)–(6.16) which are guaranteed to provide both robust stability and performance. The following existence result concerns the solvability of (6.17)–(6.21). Let n_u denote the dimension of the unstable subspace of the plant dynamics matrix A .

Theorem 7.2. Assume $n_c \geq n_u$, $R_1 > 0$, $V_1 > 0$, suppose the nominal plant, i.e., (2.1), (2.2) with $\alpha_i = 0$, $i = 1, \dots, p$, is stabilizable and detectable and, in addition, is stabilizable by means of an n_c -th-order strictly proper dynamic compensator (2.3), (2.4). Then there exist $\bar{\alpha}_1, \dots, \bar{\alpha}_p > 0$ such that if $\alpha_i \in [0, \bar{\alpha}_i]$, $i = 1, \dots, p$, then (6.17)–(6.21) have a solution $Q, P, \hat{Q}, \hat{P} \in \mathbb{R}^n$ for which (A_c, B_c, C_c) given by (6.14)–(6.16) solves the Robust Stability Problem with robust performance bound (6.22).

Proof. From Theorem 3.1 of [R1], [R2] it follows that there exists a solution to (6.17)–(6.21) which stabilizes the nominal plant. By continuity there exists a neighborhood over which robust stability with performance bound (6.22) holds. \square

Theorem 7.2 is an existence result which guarantees solvability of the sufficiency conditions over a range of parameter uncertainties. The actual range of uncertainty which can be bounded and the conservatism of the performance bound are, of course, problem dependent.

8. Specialization to Full-Order Dynamic Compensation

To draw connections with standard full-order LQG theory, we specialize the results of Sections 6 and 7 to the full-order case, i.e., $n_c = n$. As discussed in [HB], in the full-order case $G = \Gamma^{-1}$ and thus $G = \Gamma = \tau = I_n$ and $\tau_{\perp} = 0$ without loss of generality. To develop further connections with standard LQG theory assume

$$R_{12} = 0, \quad V_{12} = 0, \quad D = \Delta D = 0. \quad (8.1)$$

Since $\Delta D = 0$ we shall write $(\Delta A, \Delta B, \Delta C)$ in place of $(\Delta A, \Delta B, \Delta C, \Delta D)$. Also, for arbitrary $Q, P, \hat{Q}, \hat{P} \in \mathbb{R}^{n \times n}$ and $\alpha > 0$ define the following notation:

$$\begin{aligned} \hat{V}_{2s} &\triangleq V_2 + \sum_{i=1}^p \gamma_i C_i (Q + \hat{Q}) C_i^T, & \hat{R}_{2s} &\triangleq R_2 + \sum_{i=1}^p \gamma_i B_i^T (P + \hat{P}) B_i, \\ \hat{Q}_s &\triangleq Q C^T + \sum_{i=1}^p \gamma_i A_i (Q + \hat{Q}) C_i^T, & \hat{P}_s &\triangleq B^T P + \sum_{i=1}^p \gamma_i B_i^T (P + \hat{P}) A_i, \\ \hat{A}_Q &\triangleq A_\alpha - \hat{Q}_s \hat{V}_{2s}^{-1} C, & \hat{A}_P &\triangleq A_\alpha - B \hat{R}_{2s}^{-1} \hat{P}_s. \end{aligned}$$

Theorem 8.1. Let $n_c = n$, assume (8.1) is satisfied, and suppose there exist $Q, P, \hat{Q}, \hat{P} \in \mathbb{R}^n$ satisfying

$$0 = A_\alpha Q + Q A_\alpha^T + V_1 + \sum_{i=1}^p \gamma_i [A_i Q A_i^T + (A_i - B_i \hat{R}_{2s}^{-1} \hat{P}_s) \hat{Q} (A_i - B_i \hat{R}_{2s}^{-1} \hat{P}_s)^T] - \hat{Q}_s \hat{V}_{2s}^{-1} \hat{Q}_s^T, \quad (8.2)$$

$$0 = A_\alpha^T P + P A_\alpha + R_1 + \sum_{i=1}^p \gamma_i [A_i^T P A_i + (A_i - \hat{Q}_s \hat{V}_{2s}^{-1} C_i)^T \hat{P} (A_i - \hat{Q}_s \hat{V}_{2s}^{-1} C_i)] - \hat{P}_s^T \hat{R}_{2s}^{-1} \hat{P}_s, \quad (8.3)$$

$$0 = \hat{A}_P \hat{Q} + \hat{Q} \hat{A}_P^T + \hat{Q}_s \hat{V}_{2s}^{-1} \hat{Q}_s^T, \quad (8.4)$$

$$0 = \hat{A}_Q^T \hat{P} + \hat{P} \hat{A}_Q + \hat{P}_s^T \hat{R}_{2s}^{-1} \hat{P}_s, \quad (8.5)$$

and let (Q, A_c, B_c, C_c) be given by

$$\mathcal{Q} = \begin{bmatrix} Q + \hat{Q} & \hat{Q} \\ \hat{Q} & \hat{Q} \end{bmatrix}, \quad (8.6)$$

$$A_c = A - B \hat{R}_{2s}^{-1} \hat{P}_s - \hat{Q}_s \hat{V}_{2s}^{-1} C, \quad (8.7)$$

$$B_c = \hat{Q}_s \hat{V}_{2s}^{-1}, \quad (8.8)$$

$$C_c = -\hat{R}_{2s}^{-1} \hat{P}_s. \quad (8.9)$$

Then, $(\tilde{A} + \Delta \tilde{A}, [\tilde{V} + \Omega(\mathcal{Q}, B_c, C_c) - (\Delta \tilde{A} \mathcal{Q} + \mathcal{Q} \Delta \tilde{A}^T)^{\frac{1}{2}}])$ is stabilizable for all $(\Delta A, \Delta B, \Delta C) \in \mathcal{U}$ if and only if $\tilde{A} + \Delta \tilde{A}$ is asymptotically stable for all $(\Delta A, \Delta B, \Delta C) \in \mathcal{U}$. In this case the performance

of the closed-loop system (2.9) satisfies the bound

$$J(A_c, B_c, C_c) \leq \text{tr}[(Q + \hat{Q})R_1 + \hat{P}_e^T \hat{R}_2^{-1} R_2 \hat{R}_2^{-1} \hat{P}_e \hat{Q}]. \quad (8.10)$$

Proof. The proof follows from the reduced-order case given in Appendix B. \square

Remark 8.1. Theorem 8.1 presents sufficient conditions for robust stability and performance for full-order dynamic compensation. These sufficient conditions comprise a system of two modified Riccati equations and two Lyapunov equations coupled by the uncertainty bounds. This coupling illustrates the breakdown of regulator/estimator separation and shows that the certainty equivalence principle is no longer valid for the LQG result with real-valued structured plant parameter variations. If, however, the uncertainty terms A_i, B_i, C_i are set to zero, it can be seen that (8.4) and (8.5) drop out, while (8.2) and (8.3) reduce to the standard separated Riccati equations of LQG theory.

9. Illustrative Numerical Example

To demonstrate the above results we present an illustrative numerical example. The example chosen was originally used in [D] to illustrate the lack of a guaranteed gain margin for LQG controllers. This example was also considered in [BG1] for a preliminary robustness study. Define

$$n = 2, \quad m = \ell = p = 1,$$

$$A = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}, \quad B = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \quad C = [1 \quad 0], \quad D = 0,$$

$$A_1 = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}, \quad B_1 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \quad C_1 = [0 \quad 0], \quad D_1 = 0,$$

$$R_1 = V_1 = \begin{bmatrix} 60 & 60 \\ 60 & 60 \end{bmatrix}, \quad R_{12} = V_{12} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \quad R_2 = V_2 = 1.$$

Note that the system is open-loop unstable and becomes unstabilizable at $\sigma_1 = -1$. As can easily be seen using root locus, a strictly proper stabilizing controller must be of at least second order. Hence we consider (6.17)–(6.21) with $n_c = n$ and thus $\tau_1 = 0$. Using algorithms described in [GH],[R], controllers were obtained for $(\alpha, \alpha_1) = (.1, .1), (.4, .2)$ and $(1.6, .4)$. Figure 1 compares the guaranteed robust stability region to the actual robust stability region. Note that the design approach yields greater stability than is guaranteed a priori. This phenomenon is not surprising since even the LQG result may provide arbitrarily high levels of robustness for particular problems while failing to guarantee even minimal robustness for all problems. These results thus demonstrate

the ability of the theory to robustify the LQG result. Interestingly, the form of the actual stability region mimics the classical 6 dB downward/infinite dB upward gain margin of full-state-feedback LQR controllers. Finally, Figure 2 compares guaranteed closed-loop performance to actual closed-loop performance over the guaranteed closed-loop robust stability region. Controller gains are given in Table 1.

Acknowledgment. We wish to thank Jill M. Straehla for preparing the manuscript version of this paper and Scott W. Greeley for carrying out the numerical computations.

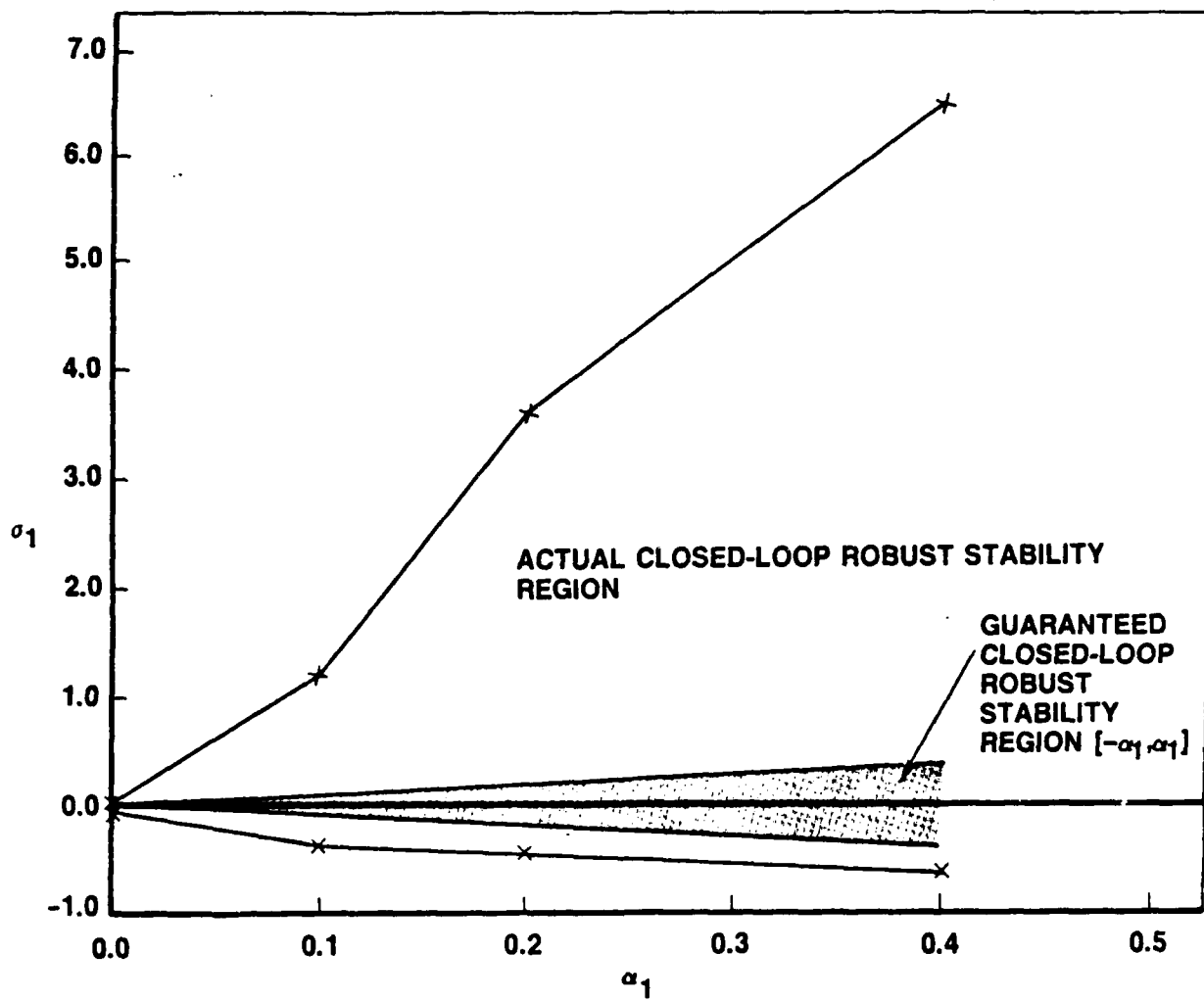


Figure 1

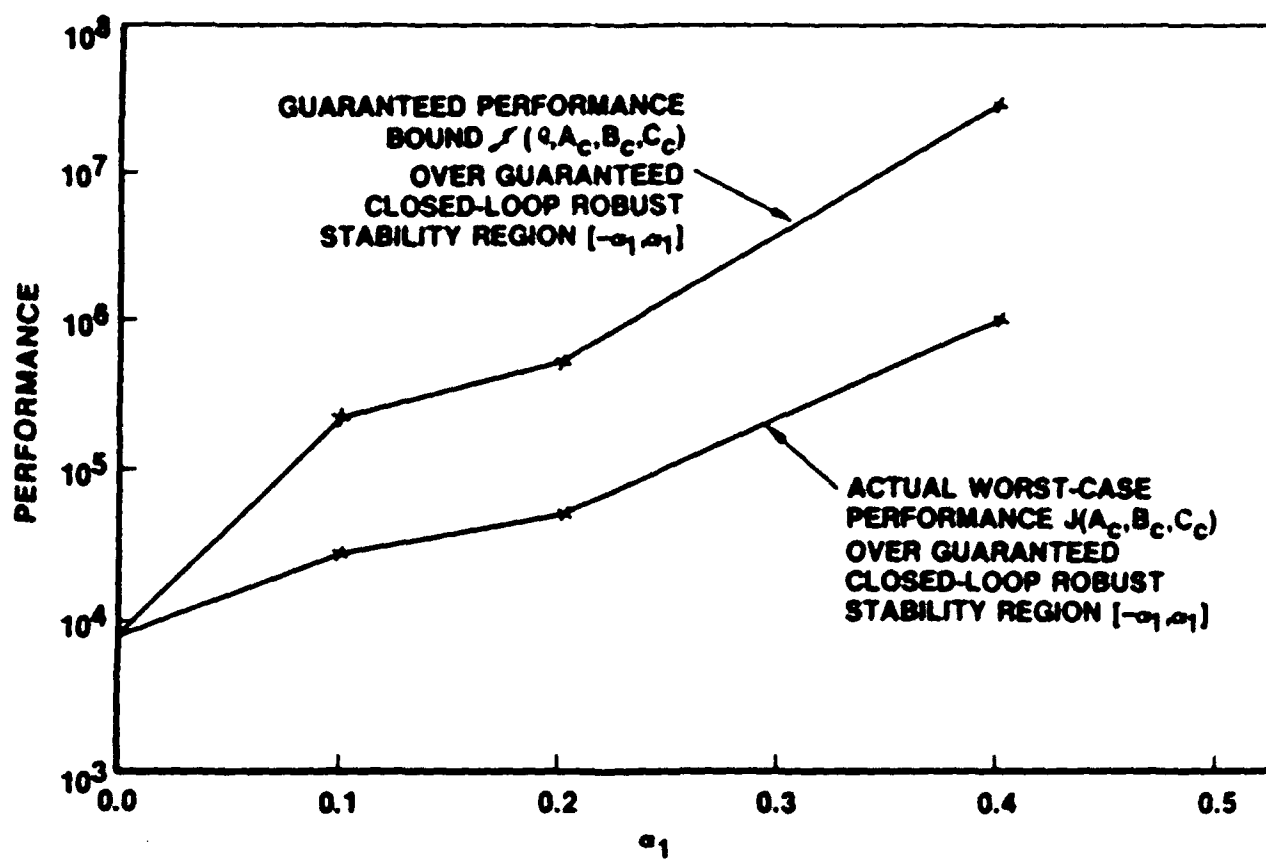


Figure 2

| (α, α_1) | A_c | B_c | C_c |
|----------------------|--|--|----------------------------|
| $(.1, .1)$ | $\begin{bmatrix} -14.917 & 1.0 \\ -85.177 & 3.9657 \end{bmatrix}$ | $\begin{bmatrix} 15.917 \\ 79.959 \end{bmatrix}$ | $[-15.2182 \quad -4.9657]$ |
| $(.4, .2)$ | $\begin{bmatrix} -17.963 & 1.0 \\ -133.65 & -4.4614 \end{bmatrix}$ | $\begin{bmatrix} 18.963 \\ 127.05 \end{bmatrix}$ | $[-6.6011 \quad -5.4614]$ |
| $(1.6, .4)$ | $\begin{bmatrix} -47.813 & 1.0 \\ -1087.3 & -6.5463 \end{bmatrix}$ | $\begin{bmatrix} 48.813 \\ 1073.5 \end{bmatrix}$ | $[-13.766 \quad -7.5463]$ |

Table 1

APPENDIX A: Proof of Lemma 6.1

- (i) Clearly $\hat{Q}\hat{P}$ and $\hat{P}^\dagger\hat{Q}\hat{P}^\dagger$ have the same nonzero eigenvalues. Since $\hat{P}^\dagger\hat{Q}\hat{P}^\dagger$ is nonnegative definite, $\hat{Q}\hat{P}$ has nonnegative eigenvalues.
- (ii) The result follows from Theorem 6.2.5 of [RM], p. 123. See also Theorem 4.3 of [G1].
- (iii) This result follows from (ii) and (6.1).
- (iv) This result follows from the definition of the group generalized inverse (see [CM]). Alternatively, let $\hat{Q}\hat{P} = \Psi D \Psi^{-1}$, where $\Psi \in D(\hat{Q}\hat{P})$, $D = \text{diag}(d_1, \dots, d_n)$. Then $(\hat{Q}\hat{P})^\# = \Psi D^\# \Psi^{-1}$, where $D^\#_{(i,i)} = 1/d_i$ if $d_i \neq 0$, and $D^\#_{(i,i)} = 0$, if $d_i = 0$, $i = 1, \dots, n$. Hence $\hat{Q}\hat{P}(\hat{Q}\hat{P})^\# = \Psi E \Psi^{-1}$ is idempotent, where E is a diagonal matrix with r ones and $n - r$ zeros on the diagonal. Clearly, (6.3) is valid.
- (v) Without loss of generality choose Ψ in the preceding argument so that $D = \text{block-diag}(\hat{D}, 0_{n-r})$, where $\hat{D} = \text{diag}(\hat{d}_1, \dots, \hat{d}_r)$, $\hat{d}_i > 0$, $i = 1, \dots, r$. Hence

$$\hat{Q}\hat{P} = \Psi \begin{bmatrix} \hat{D} & 0_{r \times (n-r)} \\ 0_{(n-r) \times r} & 0_{n-r} \end{bmatrix} \Psi^{-1},$$

and thus (6.5) holds with

$$G = [I_r \quad 0_{r \times (n-r)}] \Psi^T, \quad M = \hat{D}, \quad \Gamma = [I_r \quad 0_{r \times (n-r)}] \Psi^{-1}.$$

- (vi) Sylvester's inequality and (6.4) imply that

$$r = \text{rank } \hat{Q}\hat{P} \leq \{\text{rank } G, \text{rank } M, \text{rank } \Gamma\} \leq r,$$

which yields (6.6). The expression (6.7) for $(\hat{Q}\hat{P})^\#$ follows directly from the definition of the group generalized inverse. Furthermore, (6.2), (6.5) and (6.7) imply (6.8), while (6.5) and (6.8) imply (6.9).

- (vii) Let both (G, M, Γ) and an identically dimensioned triple $(\hat{G}, \hat{M}, \hat{\Gamma})$ satisfy (6.4). Then it is easy to verify that $\hat{G} = S^{-1}G$, $\hat{M} = SMS^{-1}$ and $\hat{\Gamma} = S\Gamma$, where $S = \hat{\Gamma}G^T$ and $S^{-1} = \Gamma\hat{G}^T$.

- (viii) It follows from (ii) that there exists $\Psi \in C(\hat{Q}, \hat{P})$ such that

$$\hat{Q} = \Psi \begin{bmatrix} D_{\hat{Q}} & 0_{r \times (n-r)} \\ 0_{(n-r) \times r} & 0_{n-r} \end{bmatrix} \Psi^T, \quad \hat{P} = \Psi^{-T} \begin{bmatrix} D_{\hat{P}} & 0_{r \times (n-r)} \\ 0_{(n-r) \times r} & 0_{n-r} \end{bmatrix} \Psi^{-1},$$

where D_Q and D_P are positive diagonal. Define

$$\hat{\Psi} = \Psi \begin{bmatrix} (D_Q D_P^{-1})^{\frac{1}{2}} & 0_{r \times (n-r)} \\ 0_{(n-r) \times r} & I_{n-r} \end{bmatrix}$$

so that

$$\hat{\Psi}^{-1} \hat{Q} \hat{\Psi}^{-T} = \hat{\Psi}^T \hat{P} \hat{\Psi} = \begin{bmatrix} (D_Q D_P)^{\frac{1}{2}} & 0_{r \times (n-r)} \\ 0_{(n-r) \times r} & 0_{n-r} \end{bmatrix}$$

and thus $\hat{\Psi} \in \mathcal{B}(\hat{Q}, \hat{P})$. Finally, (6.10) and (6.11) are immediate. \square

APPENDIX B: Proof of Theorem 6.1

To optimize (4.10) over the open set S subject to the constraint (5.5), form the Lagrangian

$$\mathcal{L}(Q, A_e, B_e, C_e, P, \lambda) \triangleq \text{tr}\{\lambda Q \tilde{R} + [\tilde{A}_e Q + Q \tilde{A}_e^T + \sum_{i=1}^p \gamma_i \tilde{A}_i Q \tilde{A}_i^T + \tilde{V}] P\}, \quad (\text{B.1})$$

where the Lagrange multipliers $\lambda \geq 0$ and $P \in \mathbb{R}^{\tilde{n} \times \tilde{n}}$ are not both zero. We thus obtain

$$\frac{\partial \mathcal{L}}{\partial Q} = \tilde{A}_e^T P + P \tilde{A}_e + \sum_{i=1}^p \gamma_i \tilde{A}_i^T P \tilde{A}_i + \lambda \tilde{R}. \quad (\text{B.2})$$

Setting $\partial \mathcal{L} / \partial Q = 0$ yields

$$0 = \tilde{A}_e^T P + P \tilde{A}_e + \sum_{i=1}^p \gamma_i \tilde{A}_i^T P \tilde{A}_i + \lambda \tilde{R} \quad (\text{B.3})$$

or, equivalently,

$$A^T \text{vec } P = -\lambda \text{vec } \tilde{R}.$$

Since A is assumed to be stable, A^T is invertible, and thus $\lambda = 0$ implies $P = 0$. Hence, it can be assumed without loss of generality that $\lambda = 1$. Furthermore, it follows from Proposition 5.2 with A, \tilde{V} replaced by A^T, \tilde{R} that P is nonnegative definite.

Now partition $\tilde{n} \times \tilde{n}$ Q, P into $n \times n, n \times n_e$, and $n_e \times n_e$ subblocks as

$$Q = \begin{bmatrix} Q_1 & Q_{12} \\ Q_{12}^T & Q_2 \end{bmatrix}, \quad P = \begin{bmatrix} P_1 & P_{12} \\ P_{12}^T & P_2 \end{bmatrix},$$

and define the positive-definite matrices

$$V_{2e} \triangleq V_2 + \sum_{i=1}^p \gamma_i [C_i Q_1 C_i^T + D_i C_e Q_2 C_e^T D_i^T], \quad R_{2e} \triangleq R_2 + \sum_{i=1}^p \gamma_i [B_i^T P_1 B_i + D_i^T B_e^T P_2 B_e D_i].$$

Thus, the stationarity conditions for A_e, B_e, C_e are given by

$$\frac{\partial \mathcal{L}}{\partial A_e} = P_{12}^T Q_{12} + P_2 Q_2 = 0, \quad (\text{B.4})$$

$$\begin{aligned} \frac{\partial \mathcal{L}}{\partial B_e} &= P_2 B_e V_{2e} + (P_{12}^T Q_1 + P_2 Q_{12}^T) C^T \\ &\quad + P_{12}^T [V_{12} + \sum_{i=1}^p \gamma_i (A_i Q_1 C_i^T + A_i Q_{12} C_e^T D_i^T)] = 0, \end{aligned} \quad (\text{B.5})$$

$$\begin{aligned} \frac{\partial \mathcal{L}}{\partial C_e} &= R_{2e} C_e Q_2 + B^T (P_1 Q_{12} + P_{12} Q_2) \\ &\quad + [R_{12}^T + \sum_{i=1}^p \gamma_i (B_i^T P_1 A_i + D_i^T B_e^T P_{12}^T A_i)] Q_{12} = 0. \end{aligned} \quad (\text{B.6})$$

Expanding (5.5) and (B.3) yields

$$0 = A_{\alpha}Q_1 + Q_1A_{\alpha}^T + BC_cQ_{12}^T + Q_{12}C_c^TB^T + \sum_{i=1}^p \gamma_i [A_iQ_1A_i^T + B_iC_cQ_{12}^TA_i^T + A_iQ_{12}C_c^TB_i^T + B_iC_cQ_2C_c^TB_i^T] + V_1, \quad (B.7)$$

$$0 = A_{\alpha}Q_{12} + Q_{12}A_{\alpha}^T + Q_{12}C_c^TD^TB_c^T + Q_1C^TB_c^T + BC_cQ_2 + \sum_{i=1}^p \gamma_i [A_iQ_1C_i^TB_c^T + A_iQ_{12}C_i^TD_i^TB_c^T] + V_{12}B_c^T, \quad (B.8)$$

$$0 = A_{\alpha\alpha}Q_2 + Q_2A_{\alpha\alpha}^T + B_cCQ_{12} + Q_{12}^TC^TB_c^T + B_cDC_cQ_2 + Q_2C_c^TD^TB_c^T + B_cV_{2\alpha}B_c^T, \quad (B.9)$$

$$0 = A_{\alpha}^TP_1 + P_1A_{\alpha} + C^TB_c^TP_{12}^T + P_{12}B_cC + \sum_{i=1}^p \gamma_i [A_i^TP_1A_i + C_i^TB_c^TP_{12}^TA_i + A_i^TP_{12}B_cC_i + C_i^TB_c^TP_2B_cC_i] + R_1, \quad (B.10)$$

$$0 = A_{\alpha}^TP_{12} + P_{12}A_{\alpha\alpha} + P_{12}B_cDC_c + P_1BC_c + C^TB_c^TP_2 + \sum_{i=1}^p \gamma_i [A_i^TP_1B_iC_c + A_i^TP_{12}B_iD_iC_c] + R_{12}C_c, \quad (B.11)$$

$$0 = A_{\alpha\alpha}^TP_2 + P_2A_{\alpha\alpha} + C_c^TB^TP_{12} + P_{12}^TBC_c + C_c^TR_{2\alpha}C_c. \quad (B.12)$$

Lemma B.1. Q_2 and P_2 are positive definite.

Proof. By a minor extension of results from [A], (B.9) can be rewritten as

$$0 = (A_{\alpha\alpha} + B_cDC_c + B_cCQ_{12}Q_2^+)Q_2 + Q_2(A_{\alpha\alpha} + B_cDC_c + B_cCQ_{12}Q_2^+)^T + B_cV_{2\alpha}B_c^T,$$

where Q_2^+ is the Moore-Penrose or Drazin generalized inverse of Q_2 . Next note that since $(A_{\alpha\alpha}, B_c)$ is controllable then, by Theorem 3.6 of [W], $(A_{\alpha\alpha} + B_cDC_c + B_cCQ_{12}Q_2^+, B_cV_{2\alpha}^{\frac{1}{2}})$ is also controllable. Now, since Q_2 and $B_cV_{2\alpha}B_c^T$ are nonnegative definite, it follows from Lemma 12.2 of [W], that Q_2 is positive definite. Using (B.12), similar arguments show that P_2 is positive definite. \square

Since $R_{2\alpha}, V_{2\alpha}, Q_2$, and P_2 are invertible, (B.4)-(B.6) can be written as

$$-P_2^{-1}P_{12}^TQ_{12}Q_2^{-1} = I_{n_s}, \quad (B.13)$$

$$B_c = -P_2^{-1} \left\{ (P_{12}^TQ_1 + P_2Q_{12}^T)C^T + P_{12}^T \left[V_{12} + \sum_{i=1}^p \gamma_i (A_iQ_1C_i^T + A_iQ_{12}C_i^TD_i^T) \right] \right\} V_{2\alpha}^{-1}, \quad (B.14)$$

$$C_c = -R_{2\alpha}^{-1} \left\{ B^T(P_1Q_{12} + P_{12}Q_2) + [R_{12}^T + \sum_{i=1}^p \gamma_i (B_i^TP_1A_i + D_i^TB_c^TP_{12}^TA_i)] Q_{12} \right\} Q_2^{-1}. \quad (B.15)$$

Now define the $n \times n$ matrices

$$\begin{aligned} Q &\triangleq Q_1 - Q_{12}Q_2^{-1}Q_{12}^T, & P &\triangleq P_1 - P_{12}P_2^{-1}P_{12}^T, \\ \hat{Q} &\triangleq Q_{12}Q_2^{-1}Q_{12}^T, & \hat{P} &\triangleq P_{12}P_2^{-1}P_{12}^T, \\ r &\triangleq -Q_{12}Q_2^{-1}P_2^{-1}P_{12}^T, \end{aligned}$$

and the $n_c \times n$, $n_c \times n_c$, and $n_c \times n$ matrices

$$G \triangleq Q_2^{-1}Q_{12}^T, \quad M \triangleq Q_2P_2, \quad \Gamma \triangleq -P_2^{-1}P_{12}^T.$$

Note that $r = G^T\Gamma$.

Clearly, Q, P, \hat{Q} , and \hat{P} are symmetric and \hat{Q} and \hat{P} are nonnegative definite. To show that Q and P are also nonnegative definite, note that Q is the upper left-hand block of the nonnegative-definite matrix $\tilde{Q}\tilde{Q}^T$, where

$$\tilde{Q} = \begin{bmatrix} I_n & -Q_{12}Q_2^{-1} \\ 0_{n_c \times n} & I_{n_c} \end{bmatrix}.$$

Similarly, P is nonnegative definite.

Next note that with the above definitions, (B.13) is equivalent to (6.5) and that (6.4) holds. Hence $r = G^T\Gamma$ is idempotent, i.e., $r^2 = r$. Furthermore, it is helpful to note the identities

$$\hat{Q} = Q_{12}G = G^TQ_{12}^T = G^TQ_2G, \quad \hat{P} = -P_{12}\Gamma = -\Gamma^TP_{12}^T = \Gamma^TP_2\Gamma, \quad (\text{B.16})$$

$$rG^T = G^T, \quad \Gamma r = \Gamma, \quad (\text{B.17})$$

$$\hat{Q} = r\hat{Q}, \quad \hat{P} = \hat{P}r, \quad (\text{B.18})$$

$$\hat{Q}\hat{P} = -Q_{12}P_{12}^T. \quad (\text{B.19})$$

Using (B.13) and Sylvester's inequality, it follows that

$$\text{rank } G = \text{rank } \Gamma = \text{rank } Q_{12} = \text{rank } P_{12} = n_c.$$

Now using (B.16) and Sylvester's inequality yields

$$n_c = \text{rank } Q_{12} + \text{rank } G - n_c \leq \text{rank } \hat{Q} \leq \text{rank } Q_{12} = n_c,$$

which implies that $\text{rank } \hat{Q} = n_c$. Similarly, $\text{rank } \hat{P} = n_c$, and $\text{rank } \hat{Q}\hat{P} = n_c$ follows from (B.19).

The components of \mathcal{Q} and \mathcal{P} can be written in terms of $Q, P, \hat{Q}, \hat{P}, G$, and Γ as

$$Q_1 = Q + \hat{Q}, \quad P_1 = P + \hat{P}, \quad (B.20)$$

$$Q_{12} = \hat{Q}\Gamma^T, \quad P_{12} = -\hat{P}G^T, \quad (B.21)$$

$$Q_2 = \Gamma\hat{Q}\Gamma^T, \quad P_2 = G\hat{P}G^T. \quad (B.22)$$

The expressions (6.13), (6.15), and (6.16) follow from the definition of \mathcal{Q} , (B.14) and (B.15). Substituting (B.20)–(B.22) into (B.7)–(B.12) yields

$$0 = A_\alpha Q + Q A_\alpha^T + V_1 + \sum_{i=1}^p \gamma_i [A_i Q A_i^T + (A_i - B_i R_{2s}^{-1} P_s) \hat{Q} (A_i - B_i R_{2s}^{-1} P_s)^T] \\ + A_P \hat{Q} + \hat{Q} A_P^T, \quad (B.23)$$

$$0 = [A_P \hat{Q} + \hat{Q} (G^T A_{\alpha\alpha} \Gamma + Q_s V_{2s}^{-1} C)^T + Q_s V_{2s}^{-1} Q_s^T] \Gamma^T, \quad (B.24)$$

$$0 = \Gamma [(G^T A_{\alpha\alpha} \Gamma + Q_s V_{2s}^{-1} C) \hat{Q} + \hat{Q} (G^T A_{\alpha\alpha} \Gamma + Q_s V_{2s}^{-1} C)^T + Q_s V_{2s}^{-1} Q_s^T] \Gamma^T, \quad (B.25)$$

$$0 = A_\alpha^T P + P A_\alpha + R_1 + \sum_{i=1}^p \gamma_i [A_i^T P A_i + (A_i - Q_s V_{2s}^{-1} C_i)^T \hat{P} (A_i - Q_s V_{2s}^{-1} C_i)] \\ + A_Q^T \hat{P} + \hat{P} A_Q, \quad (B.26)$$

$$0 = [A_Q^T \hat{P} + \hat{P} (G^T A_{\alpha\alpha} \Gamma + B R_{2s}^{-1} P_s) + P_s^T R_{2s}^{-1} P_s] G^T, \quad (B.27)$$

$$0 = G [(G^T A_{\alpha\alpha} \Gamma + B R_{2s}^{-1} P_s)^T \hat{P} + \hat{P} (G^T A_{\alpha\alpha} \Gamma + B R_{2s}^{-1} P_s) + P_s^T R_{2s}^{-1} P_s] G^T. \quad (B.28)$$

Next, computing either $\Gamma(B.24) - (B.25)$ or $G(B.27) - (B.28)$ yields (6.14). Substituting this expression for A_c into (B.23), (B.24), (B.27) and (B.28) it follows that (B.25) = $\Gamma(B.24)$ and (B.28) = $G(B.27)$. Thus, (B.25) and (B.28) are superfluous and can be omitted. Next, using (B.23) + $G^T \Gamma(B.24)G - (B.24)G - [(B.24)G]^T$ and $G^T \Gamma(B.24)G - (B.24)G - [(B.24)G]^T$ yields (6.17) and (6.19). Using (B.26) + $\Gamma^T G(B.27)\Gamma - (B.27)\Gamma - [(B.27)\Gamma]^T$ and $\Gamma^T G(B.27)\Gamma - (B.27)\Gamma - [(B.27)\Gamma]^T$ yields (6.18) and (6.20).

Finally, to prove the converse we use (6.13)–(6.21) to obtain (5.5) and (B.3)–(B.6). Let $A_c, B_c, C_c, G, \Gamma, r, Q, P, \hat{Q}, \hat{P}, \mathcal{Q}$ be as in the statement of Theorem 6.1 and define $Q_1, Q_{12}, Q_2, P_1, P_{12}, P_2$ by (B.20)–(B.22). Using (6.5), (6.15) and (6.16), it is easy to verify (B.5), (B.6). Finally, substitute the definitions of $Q, P, \hat{Q}, \hat{P}, G$, and r into (6.17)–(6.20), reverse the steps taken earlier in the proof, and use (6.13)–(6.16) along with (6.5) and (6.8)–(6.11) to obtain (5.5) and (B.3). Finally, note that

$$\mathcal{Q} = \begin{bmatrix} Q & 0_{n \times n_s} \\ 0_{n_s \times n} & 0_{n_s} \end{bmatrix} + \begin{bmatrix} I_n \\ \Gamma \end{bmatrix} \hat{Q} [I_n \quad \Gamma^T],$$

which shows that $\mathcal{Q} \geq 0$ thus verifying (4.2). \square

References

- [A] A. Albert, "Conditions for Positive and Nonnegative Definiteness in Terms of Pseudo Inverse," *SIAM J. Appl. Math.*, Vol. 17, pp. 434-440, 1969.
- [AM] B.D.O. Anderson and J.B. Moore, *Linear Optimal Control*, Prentice-Hall, Englewood Cliffs, NJ, 1970.
- [B1] B.R. Barmish, "Necessary and Sufficient Conditions for Quadratic Stabilizability of an Uncertain Linear System," *J. Optim. Thy. Appl.*, Vol. 46, pp. 399-408, 1985.
- [B2] D.S. Bernstein, "Robust Static and Dynamic Output-Feedback Stabilization: Deterministic and Stochastic Perspectives," *IEEE Trans. Autom. Contr.*, Vol. AC-32, pp. 1076-1084, 1987.
- [B3] J.W. Brewer, "Kronecker Products and Matrix Calculus in System Theory," *IEEE Trans. Circ. Sys.*, Vol. CAS-25, pp. 772-781, 1978.
- [BCL] B.R. Barmish, M. Corless, and G. Leitmann, "A New Class of Stabilizing Controllers for Uncertain Dynamical Systems," *SIAM J. Contr. Optim.*, Vol. 21, pp. 246-255, 1983.
- [BG1] D.S. Bernstein and S.W. Greeley, "Robust Controller Synthesis Using the Maximum Entropy Design Equations," *IEEE Trans. Autom. Contr.*, Vol. AC-31, pp. 362-364, 1986.
- [BG2] D.S. Bernstein and S.W. Greeley, "Robust Output-Feedback Stabilization: Deterministic and Stochastic Perspectives," *Proc. Amer. Contr. Conf.*, pp. 1818-1826, Seattle, WA, June 1986.
- [BH1] D.S. Bernstein and D.C. Hyland, "The Optimal Projection Equations for Finite-Dimensional Fixed-Order Dynamic Compensation of Infinite-Dimensional Systems," *SIAM J. Contr. Optim.*, Vol. 24, pp. 122-151, 1986.
- [BH2] D.S. Bernstein and W.M. Haddad, "The Optimal Projection Equations with Petersen-Holot Bounds: Robust Stability and Performance via Fixed-Order Dynamic Compensation for Systems with Structured Real-Valued Parameter Uncertainty," *IEEE Trans. Autom. Contr.*, Vol. AC-33, pp. 578-582, 1988.
- [BH3] D.S. Bernstein and W.M. Haddad, "Robust Stability and Performance Analysis for State Space Systems via Quadratic Lyapunov Bounds," submitted.
- [BHK] S.P. Bhattacharyya, J.W. Howze, and L.H. Keel, "Stabilization of Linear Systems with Fixed-Order Controllers," *Lin. Alg. Appl.*, Vol. 98, pp. 57-77, 1988.
- [CL] M. Corless and G. Leitmann, "Continuous State Feedback Guaranteeing Uniform Ultimate Boundedness for Uncertain Dynamical Systems," *IEEE Trans. Autom. Contr.*, Vol. AC-26, pp. 1139-1144, 1981.
- [CM] S.L. Campbell and C.D. Meyer, Jr., *Generalized Inverses of Linear Transformations*, Pitman, London, 1979.
- [CP] S.S.L. Chang and T.K.C. Peng, "Adaptive Guaranteed Cost Control of Systems with Uncertain Parameters," *IEEE Trans. Autom. Contr.*, Vol. AC-17, pp. 474-483, 1972.
- [CY] M. Cheung and S. Yurkovich, "On the Robustness of MEOP Design Versus Asymptotic LQG Synthesis," *IEEE Trans. Autom. Contr.*, Vol. 33, 1988, to appear (also, *Proc. Amer. Contr. Conf.*, pp. 24-29, Atlanta, GA, June 1988).

- [D] J.C. Doyle, "Guaranteed Margins for LQG Regulators," *IEEE Trans. Autom. Contr.*, Vol. AC-23, pp. 756-757, 1978.
- [ER] M. Eslami and D.L. Russell, "On Stability with Large Parameter Variations: Stemming from the Direct Method of Lyapunov," *IEEE Trans. Autom. Contr.*, Vol. AC-25, pp. 1231-1234, 1980.
- [G1] K. Glover, "All Optimal Hankel-Norm Approximations of Linear Multivariable Systems and Their L^∞ -Error Bounds," *Int. J. Contr.*, Vol. 39, pp. 1115-1193, 1984.
- [G2] A. Gruzen, "Robust Reduced Order Control of Flexible Structures," C. S. Draper Laboratory Report #CSDL-T-900, April 1986.
- [GB] A.R. Galimidi and B.R. Barmish, "The Constrained Lyapunov Problem and Its Application to Robust Output Feedback Stabilization," *IEEE Trans. Autom. Contr.*, Vol. AC-31, pp. 410-419, 1986.
- [GH] S.W. Greeley and D.C. Hyland, "Reduced-Order Compensation: LQG Reduction Versus Optimal Projection Using a Homotopic Continuation Method," *Proc. IEEE Conf. Dec. Contr.*, pp. 742-747, Los Angeles, CA, December 1987.
- [GV] A. Gruzen and W.E. Vander Velde, "Robust Reduced-Order Control of Flexible Structures Using the Optimal Projection/Maximum Entropy Design Methodology," *AIAA Guid. Nav. Contr. Conf.*, Williamsburg, VA, August 1978.
- [H] W.M. Haddad, *Robust Optimal Projection Control-System Synthesis*, Ph.D. Dissertation, Department of Mechanical Engineering, Florida Institute of Technology, Melbourne, FL, March 1987.
- [HB] D.C. Hyland and D.S. Bernstein, "The Optimal Projection Equations for Fixed-Order Dynamic Compensation," *IEEE Trans. Autom. Contr.*, Vol. AC-29, pp. 1034-1037, 1984.
- [KB] O.I. Kosmidou and P. Bertrand, "Robust-Controller Design for Systems with Large Parameter Variations," *Int. J. Contr.*, Vol. 45, pp. 927-938, 1987.
- [KBH] L.H. Keel, S.P. Bhattacharyya, and J.W. Howze, "Robust Control with Structured Perturbations," *IEEE Trans. Autom. Contr.*, Vol. 33, pp. 68-78, 1988.
- [KLJ] P.T. Kabamba, R.W. Longman, and S. Jian-Guo, "A Homotopy Approach to the Feedback Stabilization of Linear Systems," *J. Guid. Contr. Dyn.*, Vol. 10, pp. 422-432, 1987.
- [L] G. Leitmann, "Guaranteed Asymptotic Stability for a Class of Uncertain Linear Dynamical Systems," *J. Optim. Thy. Appl.*, Vol. 27, pp. 99-106, 1979.
- [MB] M. Mariton and R. Bertrand, "A Homotopy Algorithm for Solving Coupled Riccati Equations," *Optim. Contr. Appl. Math.*, Vol. 3, pp. 351-357, 1985.
- [PH] I.R. Petersen and C.V. Hollot, "A Riccati Equation Approach to the Stabilization of Uncertain Systems," *Automatica*, Vol. 22, pp. 397-411, 1986.
- [R1] S. Richter, "A Homotopy Algorithm for Solving the Optimal Projection Equations for Fixed-Order Dynamic Compensation: Existence, Convergence and Global Optimality," *Proc. Amer. Contr. Conf.*, pp. 1527-1531, Minneapolis, MN, June 1987.

- [R2] S. Richter, "Reduced-Order Control Design via the Optimal Projection Approach: A Homotopy Algorithm for Global Optimality," *Proc. Sixth VPI&SU Symp. Dyn. Contr. Large Str.*, L. Meirovitch, Ed., pp. 17-30, Blacksburg, VA, June 1987.
- [RM] C.R. Rao and S.K. Mitra, *Generalized Inverse of Matrices and Its Applications*, John Wiley and Sons, New York, 1971.
- [TB] J.S. Thorp and B.R. Barmish, "On Guaranteed Stability of Uncertain Linear Systems via Linear Control," *J. Optim. Thy. Appl.*, Vol. 35, pp. 559-579, 1981.
- [VW] A. Vinkler and L.J. Wood, "Multistep Guaranteed Cost Control of Linear Systems with Uncertain Parameters," *J. Guid. Contr.*, Vol. 2, pp. 449-456, 1979.
- [W] W.M. Wonham, *Linear Multivariable Control: A Geometric Approach*, Springer-Verlag, New York, 1979.

Unified optimal projection equations for simultaneous reduced-order, robust modelling, estimation and control

WASSIM M. HADDAD† and DENNIS S. BERNSTEIN‡

An optimal design problem which unifies reduced-order modelling, estimation and control problems is stated. Necessary conditions for optimality are obtained in the form of a coupled system of modified Riccati and Lyapunov equations. The results permit treatment of several new problems, such as reduced-order dynamic compensation with partially known disturbances and unified reduced-order control and estimation. Upon appropriate specialization, results obtained previously for the individual problems of reduced-order modelling, estimation and control are recovered. An additional feature is the inclusion of parameter uncertainty bounds so that the necessary conditions for an auxiliary minimization problem serve as sufficient conditions for simultaneous robust, reduced-order modelling, estimation and control.

Notation and definitions

Note. All matrices have real entries.

| | |
|--|---|
| R ; $R^{r \times s}$; R^r | real numbers; $r \times s$ real matrices; $R^{r \times 1}$ |
| I_r , $()^T$ | $r \times r$ identity matrix, transpose |
| \oplus ; \otimes | Kronecker sum; Kronecker product (Brewer 1978) |
| S^r | $r \times r$ symmetric matrices |
| N^r | $r \times r$ symmetric non-negative-definite matrices |
| P^r | $r \times r$ symmetric positive-definite matrices |
| $Z_1 \leq Z_2$ | $Z_2 - Z_1 \in N^r$, $Z_1, Z_2 \in S^r$ |
| $Z_1 < Z_2$ | $Z_2 - Z_1 \in P^r$, $Z_1, Z_2 \in S^r$ |
| asymptotically stable matrix | matrix with eigenvalues in open left half-plane |
| $n, m, \tilde{m}, l, \tilde{l}, n_c, q, p$ | positive integers |
| \tilde{n} | $n + n_c$ |
| $x, u, y, \hat{y}, x_c, y_c, y_m, \tilde{x}$ | $n, m, l, \tilde{l}, n_c, q, \tilde{n}$ -dimensional vectors |
| $A, \Delta A; B, \Delta B; C, \Delta C$ | $n \times n$ matrices; $n \times m$ matrices; $l \times n$ matrices |
| A_i, B_i, C_i | $n \times n, n \times m, l \times n$ matrices, $i = 1, \dots, p$ |
| δ_i, α_i | positive numbers, $i = 1, \dots, p$ |
| A_α | $A + \frac{1}{2} \sum_{i=1}^p \delta_i \alpha_i I_n$ |
| $A_c, B_c, B_m, C_c, C_e, C_m$ | $n_c \times n_c, n_c \times l, n_c \times \tilde{m}, m \times n_c, q \times n_c, \tilde{l} \times n_c$ matrices |
| \tilde{B}, \tilde{C} | $n \times \tilde{m}, \tilde{l} \times n$ matrices |
| A_{ca} | $A_c + \frac{1}{2} \sum_{i=1}^p \delta_i \alpha_i I_{n_c}$ |
| L | $q \times n$ matrix |

Received 9 March 1987. Revised 20 May 1987.

† Department of Mechanical Engineering, Florida Institute of Technology, Melbourne, FL 32901, U.S.A.

‡ Harris Corporation, Government Aerospace Systems Division, Melbourne, FL 32902, U.S.A.

- R estimation-error weighting in \mathbb{P}^n
 \hat{R} model-error weighting in \mathbb{P}^n
 R_1 state weighting matrix in \mathbb{R}^n
 R_2 control weighting matrix in \mathbb{P}^m
 R_{12} $n \times m$ cross weighting matrix such that $R_1 - R_{12}R_2^{-1}R_{12}^T \geq 0$
 $w(\cdot)$ m -dimensional white noise independent of $w_1(\cdot)$ and $w_2(\cdot)$
 $w_1(\cdot)$ n -dimensional white noise
 $w_2(\cdot)$ l -dimensional white noise
 V_1 intensity of $w_1(\cdot)$ in \mathbb{R}^n
 V_2 intensity of $w_2(\cdot)$ in \mathbb{P}^l
 V_{12} $n \times l$ cross intensity of $w_1(\cdot)$, $w_2(\cdot)$
 V intensity of $w(\cdot)$ in \mathbb{P}^m
 E expected value
 $\tilde{\lambda}, \Delta \tilde{\lambda}$
$$\begin{bmatrix} A & BC \\ B, C & A_1 \end{bmatrix}, \begin{bmatrix} \Delta A & \Delta BC \\ B, \Delta C & 0 \end{bmatrix}$$

 \tilde{R}
$$\begin{bmatrix} R_1 + L^T R L + \hat{C}^T \hat{R} \hat{C} & R_{12} C_1 - L^T R C_1 - \hat{C}^T \hat{R} C_{10} \\ C_1^T R_{12}^T - C_1^T R L - C_{10}^T \hat{R} \hat{C} & C_1^T R_2 C_1 + C_1^T R C_1 + C_{10}^T \hat{R} C_{10} \end{bmatrix}$$

 $\tilde{w}(t)$
$$\begin{bmatrix} w_1(t) + \hat{B} w(t) \\ B_1 w_2(t) + B_{10} w(t) \end{bmatrix}$$

 \tilde{P}
$$\begin{bmatrix} V_1 + \hat{B} V \hat{B}^T & V_{12} B_1^T + \hat{B} V B_{10}^T \\ B_1 V_{12}^T + B_{10} V \hat{B}^T & B_1 V_2 B_1^T + B_{10} V B_{10}^T \end{bmatrix}$$

1. Introduction

The problems of quadratically optimal reduced-order modelling, estimation and control have been treated in a common framework by Hyland and Bernstein (1985), Bernstein and Hyland (1985), and Hyland and Bernstein (1984), respectively. Specifically, by carrying out a judicious transformation of variables, it was shown that the necessary conditions for optimality could be cast as coupled systems of 2, 3 and 4 modified Lyapunov and Riccati equations, respectively. The coupling is via an oblique projection (i.e. idempotent matrix) which arises as a direct consequence of optimality and which determines the geometric structure of the reduced-order model, estimator, or compensator. When the order of the estimator or compensator is set equal to the order of the plant, the additional modified Lyapunov equations drop out and the remaining modified Riccati equations reduce to the standard steady-state Riccati equations of Kalman filter and LQG theory.

An immediate benefit of this formulation of the necessary conditions is clarification of the relationship between the operations of model reduction and estimator or controller design. Specifically, although the additional pair of modified Lyapunov equations arising in the reduced-order estimation and control problems are analogous to the pair of modified Lyapunov equations characterizing the optimal reduced-order model, these equations are now inextricably coupled with the modified Riccati equations characterizing the estimator and controller design. Hence, because of the coupling, this approach is distinct from LQG controller-reduction techniques (see, for

example, Liu and Anderson 1986, and Jonckheere and Silverman 1983). A comparison between the LQG reduction methods reviewed by Liu and Anderson (1986) and the optimal projection approach has been given by Greeley and Hyland (1988).

The goal of the present paper is to unify the results obtained previously for reduced-order modelling, estimation and control by deriving a single result which, upon appropriate specialization, yields the reduced-order modelling, estimation and control results as special cases. This is accomplished by defining a generalized performance functional which incorporates features of all three criteria. Thus the optimization problem involves determining a single reduced-order system which simultaneously serves as a reduced-order model, estimator and controller (or any two of these as desired). The necessary conditions now take the form of a coupled system of two modified Lyapunov equations and two modified Riccati equations which can be specialized to the necessary conditions obtained previously for the reduced-order modelling, estimation and control problems.

There are several motivations for developing a unified formulation encompassing all three results. For example, in the full-order case the certainty equivalence principle implies that the states of the optimal dynamic compensator are also optimal estimates of the plant states. This is definitely not the case for an optimal reduced-order controller in which the states may bear no resemblance to the plant states. The unified formulation of the present paper, however, expresses the desire that compensator states also provide estimates of selected plant states. Of course, except in the full-order case, such a compensator will generally be suboptimal from a strictly control point of view since the design also serves as an estimator. A similar formulation has been considered by Wilson and Kumar (1983).

Additional problems which can be handled in the unified setting involve reduced-order estimation and control in the presence of partially known plant disturbances. When measurements of disturbance components are available during real-time operation, such measurements can be used as inputs to the estimator or controller to improve performance. Note that this problem incorporates aspects of the model-reduction formulation in which the same white noise signal is injected into both the plant and the design system.

A practical motivation for the unified problem setting is convenience in developing numerical algorithms for treating different problems. In particular, a single algorithm for solving the unified optimal projection equations can readily be used for all special cases without reprogramming. (For discussions of numerical algorithms for the optimal projection equations, see Greeley and Hyland 1988, and Richter 1987.)

An additional feature of the results given herein is the treatment of parametric uncertainty in the plant matrices. By bounding the effects of parameter uncertainty on worst-case system performance, the necessary conditions for optimality effectively serve as sufficient conditions for robust stability and performance. A similar approach has been carried out by Bernstein and Haddad (1988), using structured stability radius bounds. In the present paper we use an alternative bound which corresponds to a right shift of the dynamics matrix (or equivalently, an exponential cost weighting) in conjunction with multiplicative white-noise type terms. The effect of multiplicative noise on the optimal projection equations has been developed by Bernstein and Hyland (1988). In the present paper such underlying interpretations will be suppressed since only the bound *per se* will be needed. Hence, although we use the phrase 'multiplicative white noise' for convenience in referring to the type of bound used, it should be stressed that our treatment of parameter uncertainty is wholly deterministic. (See Bernstein 1987 a, and Haddad 1987, for further background and discussion.)

2. Simultaneous reduced-order, robust modelling, estimation and control

In this section we state the 'robust performance problem' for simultaneous reduced-order modelling, estimation and control along with related notation for later use. Let $U \subset \mathbb{R}^{n \times n} \times \mathbb{R}^{n \times n} \times \mathbb{R}^{l \times n}$ denote the set of uncertain perturbations $(\Delta A, \Delta B, \Delta C)$ of the nominal system matrices A, B and C .

Robust performance problem

For fixed $n_r \leq n$, determine $(A_r, B_r, B_m, C_r, C_s, C_m)$ such that, for the augmented system consisting of the n th-order controlled and disturbed plant

$$\dot{x}(t) = (A + \Delta A)x(t) + (B + \Delta B)u(t) + \hat{B}w(t) + w_1(t), \quad t \in [0, \infty) \quad (2.1)$$

with noisy and non-noisy measurements

$$y(t) = (C + \Delta C)x(t) + w_2(t) \quad (2.2)$$

$$\hat{y}(t) = \hat{C}x(t) \quad (2.3)$$

and n_r th-order design system

$$\dot{x}_r(t) = A_r x_r(t) + B_r y(t) + B_m w(t) \quad (2.4)$$

$$u(t) = C_r x_r(t) \quad (2.5)$$

$$y_s(t) = C_s x_r(t) \quad (2.6)$$

$$y_m(t) = C_m x_r(t) \quad (2.7)$$

the performance criterion

$$J(A_r, B_r, B_m, C_r, C_s, C_m) \triangleq J_r + J_s + J_m \quad (2.8)$$

is minimized, where

$$J_r \triangleq \sup_{(\Delta A, \Delta B, \Delta C) \in U} \limsup_{t \rightarrow \infty} E[x^T(t)R_1 x(t) + 2x^T(t)R_{12}u(t) + u^T(t)R_2 u(t)] \quad (2.9)$$

$$J_s \triangleq \sup_{(\Delta A, \Delta B, \Delta C) \in U} \limsup_{t \rightarrow \infty} E[Lx(t) - y_s(t)]^T R [Lx(t) - y_s(t)] \quad (2.10)$$

$$J_m \triangleq \sup_{(\Delta A, \Delta B, \Delta C) \in U} \limsup_{t \rightarrow \infty} E[\hat{y}(t) - y_m(t)]^T \hat{R} [\hat{y}(t) - y_m(t)] \quad (2.11)$$

Remark 2.1

Suppose there are no uncertainties present, i.e. $\Delta A, \Delta B, \Delta C = 0$. By setting $L = 0$ and $\hat{C} = 0$, it follows that J_r and J_m play no role in the optimization problem when C_r and C_m are both taken to be zero. As will be seen in Theorem 6.1, this is indeed the optimal solution in this case. If, furthermore, $\hat{B} = 0$, then the reduced-order dynamic-compensation problem of Hyland and Bernstein (1984), is recovered. If, alternatively, $R_1 = 0, R_{12} = 0, B = 0, \hat{B} = 0$ and $\hat{C} = 0$ then the reduced-order state-estimation problem of Bernstein and Hyland (1985) is obtained. Finally, setting $R_1 = 0, R_{12} = 0, L = 0, V_1 = 0, B = 0$ and $C = 0$ yields the model-reduction problem considered by Hyland and Bernstein (1985).

Remark 2.2

Suppose $L = 0$ and $\hat{C} = 0$ (so that with C_r and C_s both zero J_r and J_m are ineffective) but that $\hat{B} \neq 0$. In this case, a portion of the plant disturbance, which is

assumed to be measurable during on-line operation, is being fed directly into the compensator. Hence this problem, which generalizes that of Hyland and Bernstein (1984), can be thought of as reduced-order dynamic compensation with partially known disturbances. Similarly, the case $R_1 = 0$, $R_{12} = 0$, $B = 0$ and $\hat{C} = 0$ but $\hat{B} \neq 0$ provides a generalization of Bernstein and Hyland (1985), which can be thought of as reduced-order state estimation with partially known disturbances.

For each variation $(\Delta A, \Delta B, \Delta C) \in U$, the augmented system (2.1)–(2.5) can be written as

$$\dot{\tilde{x}}(t) = (\tilde{A} + \Delta \tilde{A})\tilde{x}(t) + \tilde{w}(t), \quad t \in [0, \infty) \quad (2.12)$$

where

$$\tilde{x}(t) \triangleq [x^T(t), x_c^T(t)]^T \quad (2.13)$$

and $\tilde{w}(t)$ is white noise with intensity $\tilde{V} \in \mathbb{R}^4$.

For the 'robust performance problem' the cost can be expressed in terms of the second-moment matrix of $\tilde{x}(t)$. The following result is immediate.

Proposition 2.1

For given $(A_c, B_c, B_m, C_c, C_e, C_m)$ and $(\Delta A, \Delta B, \Delta C) \in U$ the second-moment matrix

$$\tilde{Q}_{\Delta\lambda}(t) \triangleq E[\tilde{x}(t)\tilde{x}^T(t)], \quad t \in [0, \infty) \quad (2.14)$$

satisfies

$$\dot{\tilde{Q}}_{\Delta\lambda}(t) = (\tilde{A} + \Delta \tilde{A})\tilde{Q}_{\Delta\lambda}(t) + \tilde{Q}_{\Delta\lambda}(t)(\tilde{A} + \Delta \tilde{A})^T + \tilde{V}, \quad t \in [0, \infty) \quad (2.15)$$

Furthermore,

$$J(A_c, B_c, B_m, C_c, C_e, C_m) = \sup_{(\Delta A, \Delta B, \Delta C) \in U} \limsup_{t \rightarrow \infty} \text{tr } \tilde{Q}_{\Delta\lambda}(t) \tilde{R} \quad (2.16)$$

3. Sufficient conditions for robust stability and performance

In practice, steady-state performance is only of interest when the augmented system is stable over U . The following result is immediate.

Lemma 3.1

Suppose the system (2.12) is stable for all $(\Delta A, \Delta B, \Delta C) \in U$. Then

$$J(A_c, B_c, B_m, C_c, C_e, C_m) = \sup_{(\Delta A, \Delta B, \Delta C) \in U} \text{tr } \tilde{Q}_{\Delta\lambda} \tilde{R} \quad (3.1)$$

where $\tilde{Q}_{\Delta\lambda} \in \mathbb{R}^4$ is the unique solution to

$$0 = (\tilde{A} + \Delta \tilde{A})\tilde{Q}_{\Delta\lambda} + \tilde{Q}_{\Delta\lambda}(\tilde{A} + \Delta \tilde{A})^T + \tilde{V} \quad (3.2)$$

Remark 3.1

When U is compact, 'sup' in (3.1) can be replaced by 'max'.

Since it is difficult to determine $J(A_c, B_c, B_m, C_c, C_e, C_m)$ explicitly, we shall seek upper bounds.

Theorem 3.1

Let $\Omega: \mathbb{N}^d \times \mathbb{R}^{n \times l} \times \mathbb{R}^{m \times n} \rightarrow \mathbb{S}^d$ be such that

$$\begin{aligned} \Delta \tilde{A} \mathbf{Q} + \mathbf{Q} \Delta \tilde{A}^T &\leq \Omega(\mathbf{Q}, B_c, C_c) \\ (\Delta A, \Delta B, \Delta C) &\in \mathbf{U}, \quad (\mathbf{Q}, B_c, C_c) \in \mathbb{N}^d \times \mathbb{R}^{n \times l} \times \mathbb{R}^{m \times n} \end{aligned} \quad (3.3)$$

and, for given $(A_c, B_c, B_m, C_c, C_e, C_m)$, suppose there exists $\mathbf{Q} \in \mathbb{N}^d$ satisfying

$$0 = \tilde{A} \mathbf{Q} + \mathbf{Q} \tilde{A}^T + \Omega(\mathbf{Q}, B_c, C_c) + \tilde{V} \quad (3.4)$$

Furthermore, suppose the pair $(\tilde{V}^{1/2}, \tilde{A} + \Delta \tilde{A})$ is detectable for all $(\Delta A, \Delta B, \Delta C) \in \mathbf{U}$. Then $\tilde{A} + \Delta \tilde{A}$ is asymptotically stable for all $(\Delta A, \Delta B, \Delta C) \in \mathbf{U}$,

$$\tilde{Q}_{\Delta A} \leq \mathbf{Q}, \quad (\Delta A, \Delta B, \Delta C) \in \mathbf{U} \quad (3.5)$$

and

$$J(A_c, B_c, B_m, C_c, C_e, C_m) \leq \text{tr } \mathbf{Q} \tilde{R} \quad (3.6)$$

Proof

For all $(\Delta A, \Delta B, \Delta C) \in \mathbf{U}$, (3.4) is equivalent to

$$0 = (\tilde{A} + \Delta \tilde{A}) \mathbf{Q} + \mathbf{Q} (\tilde{A} + \Delta \tilde{A})^T + \Psi(\mathbf{Q}, B_c, C_c, \Delta \tilde{A}) + \tilde{V} \quad (3.7)$$

where

$$\Psi(\mathbf{Q}, B_c, C_c, \Delta \tilde{A}) \triangleq \Omega(\mathbf{Q}, B_c, C_c) - (\Delta \tilde{A} \mathbf{Q} + \mathbf{Q} \Delta \tilde{A}^T)$$

Note that by (3.3), $\Psi(\mathbf{Q}, B_c, C_c, \Delta \tilde{A}) \geq 0$ for all $(\Delta A, \Delta B, \Delta C) \in \mathbf{U}$. Since $(\tilde{V}^{1/2}, \tilde{A} + \Delta \tilde{A})$ is detectable for all $(\Delta A, \Delta B, \Delta C) \in \mathbf{U}$, it follows from Theorem 3.6 of Wonham (1979), that $((\tilde{V} + \Psi(\mathbf{Q}, B_c, C_c, \Delta \tilde{A}))^{1/2}, \tilde{A} + \Delta \tilde{A})$ is detectable for all $(\Delta A, \Delta B, \Delta C) \in \mathbf{U}$. Hence Lemma 12.2 of Wonham (1979), implies that $\tilde{A} + \Delta \tilde{A}$ is asymptotically stable for all $(\Delta A, \Delta B, \Delta C) \in \mathbf{U}$.

Next, subtracting (3.2) from (3.7) yields

$$0 = (\tilde{A} + \Delta \tilde{A})(\mathbf{Q} - \tilde{Q}_{\Delta A}) + (\mathbf{Q} - \tilde{Q}_{\Delta A})(\tilde{A} + \Delta \tilde{A})^T + \Psi(\mathbf{Q}, B_c, C_c, \Delta \tilde{A})$$

or, equivalently (since $\tilde{A} + \Delta \tilde{A}$ is asymptotically stable),

$$\mathbf{Q} - \tilde{Q}_{\Delta A} = \int_0^\infty \exp(\tilde{A} + \Delta \tilde{A})t \Psi(\mathbf{Q}, B_c, C_c, \Delta \tilde{A}) \exp(\tilde{A} + \Delta \tilde{A})^T t dt \geq 0$$

which implies (3.5). Finally, (3.5) and (3.1) yield (3.6). \square

Remark 3.2

For the dynamic-compensation problem the result that $\tilde{A} + \Delta \tilde{A}$ is asymptotically stable for all $(\Delta A, \Delta B, \Delta C) \in \mathbf{U}$ is equivalent to robust stability of the closed-loop system. For the state-estimation and model-reduction problems, however, $\tilde{A} + \Delta \tilde{A}$ is lower block triangular (since $B = 0$) and block diagonal (since $C = 0$), respectively. Thus robust stability is equivalent to A_c stable and $A + \Delta A$ stable for all $(\Delta A, \Delta B, \Delta C) \in \mathbf{U}$.

We also note a sufficient condition for the solution \mathbf{Q} of (3.4) to be positive definite.

Proposition 3.1

Let Ω be as in Theorem 3.1, let $(A_c, B_c, B_m, C_c, C_e, C_m)$ be given, and suppose there exists $Q \in \mathbb{R}^n$ satisfying (3.4). If $(\tilde{V}^{1/2}, \tilde{A} + \Delta\tilde{A})$ is observable for some $(\Delta A, \Delta B, \Delta C) \in U$, then Q is positive definite.

Proof

If $(\tilde{V}^{1/2}, \tilde{A} + \Delta\tilde{A})$ is observable for some $(\Delta A, \Delta B, \Delta C) \in U$, then, by Theorem 3.6 of Wonham (1979), $((\tilde{V} + \Psi(Q, B_c, C_c, \Delta\tilde{A}))^{1/2}, \tilde{A} + \Delta\tilde{A})$ is also observable for the same $(\Delta A, \Delta B, \Delta C) \in U$. It thus follows from (3.7) and Lemma 12.2 of Wonham (1979), that Q is positive definite. \square

Remark 3.3

If \tilde{V} is positive definite then the detectability and observability hypotheses of Theorem 3.1 and Proposition 3.1 are automatically satisfied.

Remark 3.4

Theorem 3.1 can be strengthened by noting that the detectability assumption is, in a sense, superfluous. To see this, first note that robust stability concerns only the undisturbed system while \tilde{V} involves the disturbance noise. Hence robust stability is guaranteed by the existence of a solution $Q \in \mathbb{R}^n$ satisfying (3.4) with \tilde{V} replaced by αI_n for some $\alpha > 0$. For this replacement detectability is automatic (see previous remark). For robust performance, however, Q in (3.5) must be obtained from (3.4).

4. Uncertainty structure and right shift multiplicative white noise bound

The uncertainty set U is assumed to be of the form

$$U = \left\{ (\Delta A, \Delta B, \Delta C) \in \mathbb{R}^{n \times n} \times \mathbb{R}^{n \times m} \times \mathbb{R}^{l \times n} : \Delta A = \sum_{i=1}^p \sigma_i A_i, \right. \\ \left. \Delta B = \sum_{i=1}^p \sigma_i B_i, \quad \Delta C = \sum_{i=1}^p \sigma_i C_i, \quad |\sigma_i| \leq \delta_i, \quad i = 1, \dots, p \right\} \quad (4.1)$$

where, for $i = 1, \dots, p$: $A_i \in \mathbb{R}^{n \times n}$, $B_i \in \mathbb{R}^{n \times m}$ and $C_i \in \mathbb{R}^{l \times n}$ are fixed matrices denoting the structure of the parametric uncertainty; δ_i is a given uncertainty bound; and σ_i is an uncertain real parameter. The closed-loop system thus has structured uncertainty of the form

$$\Delta\tilde{A} = \sum_{i=1}^p \sigma_i \tilde{A}_i$$

where

$$\tilde{A}_i \triangleq \begin{bmatrix} A_i & B_i C_c \\ B_c C_i & 0 \end{bmatrix}, \quad i = 1, \dots, p \quad (4.2)$$

To obtain an explicit gain expression for $(A_c, B_c, B_m, C_c, C_e, C_m)$ we require that

$$[B_i \neq 0 \Rightarrow C_i = 0], \quad i = 1, \dots, p \quad (4.3)$$

That is, for each $i \in \{1, \dots, p\}$ either B_i or C_i is zero. Of course, both $B_i = 0$ and $C_i = 0$ are possible for a given i , and there are no restrictions on A_i .

Given the structure of \mathbf{U} defined by (4.1) we can define the bound satisfying (3.3).

Proposition 4.1

Let $\alpha_1, \dots, \alpha_p$ be arbitrary positive scalars. Then the function

$$\Omega(\mathbf{Q}, B_c, C_c) = \sum_{i=1}^p \delta_i (\alpha_i \mathbf{Q} + \alpha_i^{-1} \tilde{A}_i \mathbf{Q} \tilde{A}_i^T) \quad (4.4)$$

satisfies (3.3) with \mathbf{U} given by (4.1).

Proof

Note that

$$\begin{aligned} 0 &\leq [\sigma_i(\alpha_i/\delta_i)^{1/2} I_n - (\delta_i/\alpha_i)^{1/2} \tilde{A}_i] \mathbf{Q} [\sigma_i(\alpha_i/\delta_i)^{1/2} I_n - (\delta_i/\alpha_i)^{1/2} \tilde{A}_i]^T \\ &= \sigma_i^2(\alpha_i/\delta_i) \mathbf{Q} + (\delta_i/\alpha_i) \tilde{A}_i \mathbf{Q} \tilde{A}_i^T - \sigma_i(\tilde{A}_i \mathbf{Q} + \mathbf{Q} \tilde{A}_i^T) \end{aligned}$$

which, since $\sigma_i^2 \leq \delta_i^2$, implies (3.3). \square

5. Auxiliary minimization problem

Rather than minimize the actual cost (2.8), we shall consider the upper bound (3.6). This leads to the following problem.

Auxiliary minimization problem

Determine $(A_c, B_c, B_m, C_c, C_e, C_m)$ and $\mathbf{Q} \in \hat{\mathbf{N}}^n$ which minimize

$$J(A_c, B_c, B_m, C_c, C_e, C_m, \mathbf{Q}) \triangleq \text{tr } \mathbf{Q} \tilde{R} \quad (5.1)$$

subject to

$$0 = \tilde{A} \mathbf{Q} + \mathbf{Q} \tilde{A}^T + \sum_{i=1}^p \delta_i (\alpha_i \mathbf{Q} + \alpha_i^{-1} \tilde{A}_i \mathbf{Q} \tilde{A}_i^T) + \tilde{V} \quad (5.2)$$

and

$$(\tilde{V}^{1/2}, \tilde{A} + \Delta \tilde{A}) \text{ is detectable, } (\Delta A, \Delta B, \Delta C) \in \mathbf{U} \quad (5.3)$$

Proposition 5.1

If $(A_c, B_c, B_m, C_c, C_e, C_m, \mathbf{Q})$ is admissible, i.e. $(A_c, B_c, B_m, C_c, C_e, C_m, \mathbf{Q})$ satisfies (5.2) and (5.3), then $\tilde{A} + \Delta \tilde{A}$ is asymptotically stable for all $(\Delta A, \Delta B, \Delta C) \in \mathbf{U}$ and

$$J(A_c, B_c, B_m, C_c, C_e, C_m) \leq J(A_c, B_c, B_m, C_c, C_e, C_m, \mathbf{Q}) \quad (5.4)$$

Proof

With Ω given by (4.4), Proposition 4.1 implies that (3.3) is satisfied. Furthermore, admissibility implies that (3.4) has a solution $\mathbf{Q} \in \hat{\mathbf{N}}^n$. Hence, with (5.3), the hypotheses of Theorem 3.1 are satisfied so that robust stability with the performance bound (3.6) is guaranteed. Note that with the definition (5.1), (5.4) is merely a restatement of (3.6). \square

6. Necessary conditions for the auxiliary minimization problem

The derivation of the necessary conditions for the 'auxiliary minimization problem' is based upon the Fritz John form of the Lagrange multiplier theorem.

Rigorous application of this technique requires additional technical assumptions. Specifically, we further restrict $(A_c, B_c, B_m, C_c, C_e, C_m)$ to the set

$$\mathbf{S} \triangleq \{(A_c, B_c, B_m, C_c, C_e, C_m, \mathbf{Q}) : \mathbf{Q} \in \hat{\mathbf{P}}, \bar{\mathbf{A}} \text{ is asymptotically stable, and } (A_c, B_c, C_c) \text{ is minimal}\}$$

where

$$\bar{\mathbf{A}} \triangleq \left(\bar{\mathbf{A}} + \frac{1}{2} \sum_{i=1}^p \delta_i \alpha_i I_n \right) \oplus \left(\bar{\mathbf{A}} + \frac{1}{2} \sum_{i=1}^p \delta_i \alpha_i I_n \right) + \sum_{i=1}^p \gamma_i \bar{\mathbf{A}}_i \otimes \bar{\mathbf{A}}_i$$

with, for convenience,

$$\gamma_i \triangleq \delta_i / \alpha_i$$

The following observation assures us that we can apply Lagrange multipliers over an open constraint set.

Proposition 6.1

The set \mathbf{S} is open.

Proof

It need only be noted that \mathbf{S} is the intersection of three open sets. \square

Remark 6.1

The constraint $(A_c, B_c, B_m, C_c, C_e, C_m, \mathbf{Q}) \in \mathbf{S}$ is not required for either robust stability or robust performance. As can be seen from the proof of Theorem 6.1 in the Appendix, the set \mathbf{S} constitutes sufficient conditions under which the Lagrange multiplier technique is applicable to the 'auxiliary minimization problem'. Specifically, asymptotic stability of $\bar{\mathbf{A}}$ serves as a normality condition which further implies that the dual \mathbf{P} of \mathbf{Q} is non-negative definite. Furthermore, (A_c, B_c, C_c) minimal is a nondegeneracy condition which implies that the lower right $n_c \times n_c$ subblocks of \mathbf{Q} and \mathbf{P} are positive definite. It is extremely important to emphasize that Proposition 5.1 implies that it is *not* necessary for guaranteed robust stability and performance that an admissible quadruple, obtained by solving the necessary conditions, actually be shown to be an element of \mathbf{S} .

For arbitrary $\mathbf{Q}, \mathbf{P}, \hat{\mathbf{Q}}, \hat{\mathbf{P}} \in \mathbf{R}^{n \times n}$ define the following notation:

$$\begin{aligned} R_{2s} &\triangleq R_2 + \sum_{i=1}^p \gamma_i B_i^T (P + \hat{P}) B_i, & V_{2s} &\triangleq V_2 + \sum_{i=1}^p \gamma_i C_i (Q + \hat{Q}) C_i^T \\ Q_s &\triangleq Q C^T + V_{12} + \sum_{i=1}^p \gamma_i A_i (Q + \hat{Q}) C_i^T, & P_s &\triangleq B^T P + R_{12}^T + \sum_{i=1}^p \gamma_i B_i^T (P + \hat{P}) A_i \\ A_Q &\triangleq A_s - Q_s V_{2s}^{-1} C, & A_P &\triangleq A_s - B R_{2s}^{-1} P_s \end{aligned}$$

The following factorization lemma is needed for the statement of the main result.

Lemma 6.1

If $\hat{\mathbf{Q}}, \hat{\mathbf{P}} \in \hat{\mathbf{N}}$ then $\hat{\mathbf{Q}}\hat{\mathbf{P}}$ is diagonalizable with non-negative eigenvalues. If, in addition, $\text{rank } \hat{\mathbf{Q}}\hat{\mathbf{P}} = n_c$, then there exist $n_c \times n$ G, Γ and $n_c \times n_c$ invertible M such that

$$\hat{\mathbf{Q}}\hat{\mathbf{P}} = G^T M \Gamma \quad (6.1)$$

$$\Gamma G^T = I_{n_c} \quad (6.2)$$

Furthermore, G , M and Γ are unique except for a change of basis in \mathbb{R}^{n_c} .

Proof

The result is an immediate consequence of Rao and Mitra (1971), Theorem 6.2.5, p. 123. \square

A triple (G, M, Γ) satisfying (6.1) and (6.2) will be called a *projective factorization* of $\hat{Q}\hat{P}$. Since $\hat{Q}\hat{P}$ is diagonalizable it has a group generalized inverse $(\hat{Q}\hat{P})^+ = G^T M^{-1} \Gamma$ and

$$\tau \triangleq \hat{Q}\hat{P}(\hat{Q}\hat{P})^+ = G^T \Gamma$$

is an oblique projection. Furthermore, define the complementary projection

$$\tau_\perp \triangleq I_n - \tau.$$

Theorem 6.1

Suppose $(A_c, B_c, B_m, C_c, C_e, C_m, Q) \in \mathbf{S}$ solves the 'auxiliary minimization problem'. Then there exist $Q, P, \hat{Q}, \hat{P} \in \hat{\mathbf{N}}^n$ such that, for some projective factorization (G, M, Γ) of $\hat{Q}\hat{P}$, $A_c, B_c, B_m, C_c, C_e, C_m$ and Q are given by

$$A_c = \Gamma(A - BR_{2s}^{-1}P_s - Q_s V_{2s}^{-1}C)G^T \quad (6.3)$$

$$B_c = \Gamma Q_s V_{2s}^{-1} \quad (6.4)$$

$$B_m = \Gamma \hat{B} \quad (6.5)$$

$$C_c = -R_{2s}^{-1}P_s G^T \quad (6.6)$$

$$C_e = LG^T \quad (6.7)$$

$$C_m = \hat{C}G^T \quad (6.8)$$

$$Q = \begin{bmatrix} Q + \hat{Q} & \hat{Q}\Gamma^T \\ \Gamma\hat{Q} & \Gamma\hat{Q}\Gamma^T \end{bmatrix} \quad (6.9)$$

and such that Q, P, \hat{Q} and \hat{P} satisfy

$$0 = A_s Q + Q A_s^T + V_1 + \sum_{i=1}^p \gamma_i [A_i Q A_i^T + (A_i - B_i R_{2s}^{-1} P_s) \hat{Q} (A_i - B_i R_{2s}^{-1} P_s)^T] \\ - Q_s V_{2s}^{-1} Q_s^T + \tau_\perp [Q_s V_{2s}^{-1} Q_s^T + \hat{B} V \hat{B}^T] \tau_\perp^T \quad (6.10)$$

$$0 = A_s^T P + P A_s + R_1 + \sum_{i=1}^p \gamma_i [A_i^T P A_i + (A_i - Q_s V_{2s}^{-1} C_i)^T \hat{P} (A_i - Q_s V_{2s}^{-1} C_i) \\ - P_s^T R_{2s}^{-1} P_s + \tau_\perp^T [P_s^T R_{2s}^{-1} P_s + L^T R L + \hat{C}^T \hat{R} \hat{C}] \tau_\perp \quad (6.11)$$

$$0 = A_p \hat{Q} + \hat{Q} A_p^T + Q_s V_{2s}^{-1} Q_s^T + \hat{B} V \hat{B}^T - \tau_\perp [Q_s V_{2s}^{-1} Q_s^T + \hat{B} V \hat{B}^T] \tau_\perp^T \quad (6.12)$$

$$0 = A_Q^T \hat{P} + \hat{P} A_Q + P_s^T R_{2s}^{-1} P_s + L^T R L + \hat{C}^T \hat{R} \hat{C} - \tau_\perp^T [P_s^T R_{2s}^{-1} P_s + L^T R L + \hat{C}^T \hat{R} \hat{C}] \tau_\perp \quad (6.13)$$

$$\text{rank } \hat{Q} = \text{rank } \hat{P} = \text{rank } \hat{Q}\hat{P} = n_c \quad (6.14)$$

Proof

For the proof see the Appendix.

Remark 6.2

As in Remark 2.1 suppose $\Delta A, \Delta B, \Delta C = 0$. By setting $L = 0, \hat{B} = 0$ and $C = 0$, (6.10)–(6.13) specialize to the optimal projection equations (2.18)–(2.21) derived by Hyland and Bernstein (1984), with the added features of correlated plant/measurement noise (V_{12}) and cross weighting (R_{12}). If $R_1 = 0, R_{12} = 0, B = 0, \hat{B} = 0$ and $\hat{C} = 0$ then, since $P_s = 0$, (6.11) drops out and the remaining equations (6.10), (6.12) and (6.13) specialize to (2.10)–(2.12) of Bernstein and Hyland (1985). Finally, if $R_1 = 0, R_{12} = 0, L = 0, V_1 = 0, B = 0$ and $C = 0$, then, since $Q_s = P_s = 0$, (6.10) and (6.11) drop out and the remaining equations (6.12) and (6.13) specialize to (2.21) and (2.22) of Hyland and Bernstein (1985).

Remark 6.3

A more restrictive formulation for unified modelling, estimation and control is to require $C_c = C_e = C_m$ so that $u = y_e = y_m$. In this case the three outputs of the design system (2.4)–(2.7) are replaced by a single output. Again, the necessary conditions involve a system of four coupled matrix equations similar to (6.10)–(6.13) which specialize to previously known results. Since this formulation requires $m = q = l$, it appears to be less useful than the three-output formulation.

7. Sufficient conditions for robust stability and performance

The main result guaranteeing robust stability and performance for the unified problem can now be stated.

Theorem 7.1

Suppose there exist $Q, P, \hat{Q}, \hat{P} \in \mathbb{N}^n$ satisfying (6.10)–(6.14) and assume $(\tilde{V}^{1/2}, \tilde{A} + \Delta\tilde{A})$ is detectable for all $(\Delta A, \Delta B, \Delta C) \in \mathbf{U}$ with $A_c, B_c, B_m, C_c, C_e, C_m$ given by (6.3)–(6.18). Then $\tilde{A} + \Delta\tilde{A}$ is asymptotically stable for all $(\Delta A, \Delta B, \Delta C) \in \mathbf{U}$ and the closed-loop system satisfies the performance bound

$$J(A_c, B_c, B_m, C_c, C_e, C_m) \leq \text{tr}[(Q + \hat{Q})R_1 + P_s^T R_{2s}^{-1} R_2 R_{2s}^{-1} P_s \hat{Q} - 2R_{12} R_{2s}^{-1} P_s \hat{Q} + QL^T RL + \hat{C}^T \hat{R} \hat{C}(W_c - \hat{Q})] \quad (7.1)$$

where the controllability gramian W_c satisfies

$$0 = AW_c + W_c A^T + \hat{B}V\hat{B}^T \quad (7.2)$$

Proof

Theorem 7.1 implies \mathbf{Q} given by (6.9) satisfies (5.2). With the detectability assumption the result follows from Proposition 5.1. \square

8. Directions for further research

Several generalizations remain to be explored. These include:

- (a) permit $w(\cdot)$ to be correlated with $w_1(\cdot)$ and $w_2(\cdot)$;

(b) replace (2.2) with

$$y(t) = (C + \Delta C)x(t) + (D + \Delta D)u(t) + w_2(t) \quad (8.1)$$

(c) replace (2.4) with

$$\dot{x}_c(t) = A_c x_c(t) + B_c y(t) + B_m w(t) + w_3(t) \quad (8.2)$$

(d) replace (2.5) with

$$u(t) = C_c x_c(t) + D_c y(t) \quad (8.3)$$

The extension (8.3) has been studied by Bernstein (1987 b), for control and by Haddad and Bernstein (1987), for estimation.

ACKNOWLEDGMENT

The authors are indebted to Scott W. Greeley for bringing to their attention the reduced-order state estimation problem with partially known disturbances which motivated this paper. We also wish to thank Jill Straehla for excellent typing of the original manuscript.

Dr Bernstein was supported in part by the Air Force Office of Scientific Research under Contract AFOSR F49620-86-C-0002.

Appendix

Proof of Theorem 6.1

Partition $\tilde{n} \times \tilde{n}$ \mathbf{Q} , \mathbf{P} into $n \times n$, $n \times n_c$, and $n_c \times n_c$ sub-blocks as

$$\mathbf{Q} = \begin{bmatrix} Q_1 & Q_{12} \\ Q_{12}^T & Q_2 \end{bmatrix}, \quad \mathbf{P} = \begin{bmatrix} P_1 & P_{12} \\ P_{12}^T & P_2 \end{bmatrix}$$

and define the $n \times n$ matrices

$$Q \triangleq Q_1 - Q_{12} Q_2^{-1} Q_{12}^T, \quad P \triangleq P_1 - P_{12} P_2^{-1} P_{12}^T \\ \hat{Q} \triangleq Q_{12} Q_2^{-1} Q_{12}^T, \quad \hat{P} \triangleq P_{12} P_2^{-1} P_{12}^T$$

and the $n_c \times n$, $n_c \times n_c$, $n_c \times n$ matrices

$$G \triangleq Q_2^{-1} Q_{12}^T, \quad M \triangleq Q_2 P_2, \quad \Gamma \triangleq -P_2^{-1} P_{12}^T$$

The existence of Q_2^{-1} and P_2^{-1} is shown below.

Clearly, Q , P , \hat{Q} and \hat{P} are symmetric and \hat{Q} and \hat{P} are non-negative definite. To show that Q and P are non-negative definite, note that Q is the upper left-hand block of the non-negative-definite matrix $Q^* \mathbf{Q} Q^{*T}$, where

$$Q^* \triangleq \begin{bmatrix} I_n & -Q_{12} Q_2^{-1} \\ 0 & -I_{n_c} \end{bmatrix}$$

Similarly, P is non-negative definite.

To optimize (5.1) over the open set \mathbf{S}' , where

$$\mathbf{S}' \triangleq \{(A_c, B_c, B_m, C_c, C_e, C_m, \mathbf{Q}) \in \mathbf{S}: (5.3) \text{ is satisfied}\}$$

and subject to the constraint (5.2), form the lagrangian

$$L(A_c, B_c, B_m, C_c, C_e, C_m, Q, P, \lambda) \\ \triangleq \text{tr} \left[\lambda Q \tilde{R} + \left(\tilde{A} Q + Q \tilde{A}^T + \sum_{i=1}^p \delta_i (\alpha_i Q + \alpha_i^{-1} \tilde{A}_i Q \tilde{A}_i^T) + \tilde{V} \right) P \right]$$

where the Lagrange multipliers $\lambda \geq 0$ and $P \in \mathbb{P}^{n \times n}$ are not both zero. We thus obtain

$$\frac{\partial L}{\partial Q} = \tilde{A}^T P + P \tilde{A} + \sum_{i=1}^p \delta_i (\alpha_i P + \alpha_i^{-1} \tilde{A}_i^T P \tilde{A}_i) + \lambda \tilde{R}$$

Setting $\partial L / \partial Q = 0$ yields

$$\tilde{A}^T \text{vec } P = -\lambda \text{vec } \tilde{R}$$

where 'vec' is the column-stacking operation (see Brewer 1978). Since \tilde{A} is assumed to be stable and thus invertible, $\lambda = 0$ implies $P = 0$. Hence, it can be assumed without loss of generality that $\lambda = 1$.

Furthermore, the stability of \tilde{A} implies that P is non-negative definite. The stationarity conditions are given by

$$\frac{\partial L}{\partial P} = \tilde{A} Q + Q \tilde{A}^T + \sum_{i=1}^p [\delta_i \alpha_i Q + \gamma_i \tilde{A}_i Q \tilde{A}_i^T] + \tilde{V} = 0 \quad (\text{A } 1)$$

$$\frac{\partial L}{\partial Q} = \tilde{A}^T P + P \tilde{A} + \sum_{i=1}^p [\delta_i \alpha_i P + \gamma_i \tilde{A}_i^T P \tilde{A}_i] + \tilde{R} = 0 \quad (\text{A } 2)$$

$$\frac{\partial L}{\partial A_c} = P_{12}^T Q_{12} + P_2 Q_2 = 0 \quad (\text{A } 3)$$

$$\frac{\partial L}{\partial B_c} = P_2 B_c V_{2s} + (P_{12}^T Q_1 + P_2 Q_{12}^T) C + P_{12}^T \left(V_{12} + \sum_{i=1}^p \gamma_i A_i Q_1 C_i^T \right) = 0 \quad (\text{A } 4)$$

$$\frac{\partial L}{\partial B_m} = P_{12}^T \hat{B} V + P_2 B_m V = 0 \quad (\text{A } 5)$$

$$\frac{\partial L}{\partial C_c} = R_{2s} C_c Q_2 + B^T (P_1 Q_{12} + P_{12} Q_2) + \left(R_{12}^T + \sum_{i=1}^p \gamma_i B_i^T P_1 A_i \right) Q_{12} = 0 \quad (\text{A } 6)$$

$$\frac{\partial L}{\partial C_e} = -R L Q_{12} + R C_e Q_{12} = 0 \quad (\text{A } 7)$$

$$\frac{\partial L}{\partial C_m} = -\hat{R} \hat{C} Q_{12} + \hat{R} C_m Q_2 = 0 \quad (\text{A } 8)$$

Expanding (A 1) and (A 2) yields

$$0 = A_s Q_1 + B C_c Q_{12}^T + Q_1 A_s^T + Q_{12} C_c^T B^T + \sum_{i=1}^p \gamma_i \\ \times [A_i Q_1 A_i^T + B_i C_c Q_{12}^T A_i^T + A_i Q_{12} C_c^T B_i^T + B_i C_c Q_2 C_i^T B_i^T] + V_1 + \hat{B} V \hat{B}^T \quad (\text{A } 9)$$

$$0 = A_s Q_{12} + Q_{12} A_{cs}^T + B C_c Q_2 + Q_1 C^T B_c^T + \sum_{i=1}^p \gamma_i A_i Q_1 C_i^T B_c^T \\ + V_{12} B_c^T + \hat{B} V B_m^T \quad (\text{A } 10)$$

$$0 = B_c C Q_{12} + A_{cs} Q_2 + Q_{12}^T C^T B_c^T + Q_2 A_{cs}^T + B_c V_{2s} B_c^T + B_m V B_m^T \quad (\text{A } 11)$$

$$0 = P_1 A_a + P_{12} B_c C + A_a^T P_1 + C^T B_c^T P_{12}^T + \sum_{i=1}^p \gamma_i \times [A_i^T P_1 A_i + C_i^T B_c^T P_{12}^T A_i + A_i^T P_{12} B_c C_i + C_i^T B_c^T P_2 B_c C_i] + R_1 + L^T R L + \hat{C}^T \hat{R} \hat{C} \quad (\text{A } 12)$$

$$0 = P_1 B C_c + P_{12} A_{ca} + A^T P_{12} + C^T B_c^T P_2 + \sum_{i=1}^p \gamma_i A_i^T P_1 B_i C_c + R_{12} C_c - L^T R C_m - \hat{C}^T R \hat{C}_m \quad (\text{A } 13)$$

$$0 = A_{ca}^T P_2 + P_2 A_{ca} + C_c^T B^T P_{12} + P_{12}^T B C_c + C_c^T R_{2s} C_c + C_e^T R C_e + C_m^T \hat{R} C_m \quad (\text{A } 14)$$

Lemma A 1

Q_2 and P_2 are positive definite.

Proof

By a minor extension of the results from Albert (1969), (A 11) can be rewritten as

$$0 = (A_{ca} + B_c C Q_{12} Q_2^+) Q_2 + Q_2 (A_{ca} + B_c C Q_{12} Q_2^+)^T + \Lambda$$

where $\Lambda = B_c V_{2s} B_c^T + B_m V B_m^T$ and Q_2^+ is the Moore-Penrose or Drazin generalized inverse of Q_2 . Next note that since (A_c, B_c) is controllable then, by Theorem 3.6 of Wonham (1979), $(A_{ca} + B_c C Q_{12} Q_2^+, \Lambda^{1/2})$ is also controllable. Now, since Q_2 and Λ are non-negative definite, it follows from Lemma 12.2 of Wonham (1979), that Q_2 is positive definite. Using (A 14), similar arguments show P_2 is positive definite. \square

Since \hat{R} , R , R_{2s} , V , V_{2s} , Q_2 and P_2 are invertible, (A 3)–(A 8) can be written as

$$-P_2^{-1} P_{12}^T Q_{12} Q_2^{-1} = I_n \quad (\text{A } 15)$$

$$B_c = -P_2^{-1} \left[(P_{12}^T Q_1 + P_2 Q_{12}^T) C + P_{12}^T \left(V_{12} + \sum_{i=1}^p \gamma_i A_i Q_1 C_i^T \right) \right] V_{2s}^{-1} \quad (\text{A } 16)$$

$$B_m = -P_2^{-1} P_{12} \hat{B} \quad (\text{A } 17)$$

$$C_c = -R_{2s}^{-1} \left[B^T (P_1 Q_{12} + P_{12} Q_2) + \left(R_{12}^T + \sum_{i=1}^p \gamma_i B_i^T P_1 A_i \right) Q_{12} \right] Q_2^{-1} \quad (\text{A } 18)$$

$$C_e = L Q_{12} Q_2^{-1} \quad (\text{A } 19)$$

$$C_m = \hat{C} Q_{12} Q_2^{-1} \quad (\text{A } 20)$$

Note that because of (A 15), (6.1) and (6.2) hold. Since Q_2 and P_2 are positive definite and

$$Q_2 P_2 = P_2^{-1/2} (P_2^{1/2} Q_2 P_2^{1/2}) P_2^{1/2}$$

M is diagonalizable with positive eigenvalues. It is helpful to note the identities

$$\hat{Q} = Q_{12} G = G^T Q_{12}^T = G^T Q_2 G, \quad \hat{P} = -P_{12} \Gamma = -\Gamma^T P_{12}^T = \Gamma^T P_2 \Gamma \quad (\text{A } 21)$$

$$\tau G^T = G^T, \quad \Gamma \tau = \Gamma \quad (\text{A } 22)$$

$$\hat{Q} = \tau \hat{Q}, \quad \hat{P} = \hat{P} \tau \quad (\text{A } 23)$$

$$\hat{Q} \hat{P} = -Q_{12} P_{12}^T \quad (\text{A } 24)$$

Using (6.2) and Sylvester's inequality, it follows that $\text{rank } G = \text{rank } \Gamma = \text{rank } Q_{12} = \text{rank } P_{12} = n_c$ which in turn imply (6.14).

The components of Q and P can be written in terms of $Q, P, \hat{Q}, \hat{P}, G$ and Γ as

$$Q_1 = Q + \hat{Q}, \quad P_1 = P + \hat{P} \quad (\text{A } 25)$$

$$Q_{12} = \hat{Q}\Gamma^T, \quad P_{12} = -\hat{P}G^T \quad (\text{A } 26)$$

$$Q_2 = \Gamma\hat{Q}\Gamma^T, \quad P_2 = G\hat{P}G^T \quad (\text{A } 27)$$

The gain expressions (6.3)–(6.8) and (6.9) follow from (A 16)–(A 20) and the definition of \mathbf{Q} . Substituting (A 25)–(A 27) into (A 9)–(A 14) yields

$$0 = A_c Q + Q A_c^T + V_1 + \hat{B} V \hat{B}^T + \sum_{i=1}^p \gamma_i \times [A_i Q A_i^T + (A_i - B_i R_{2s}^{-1} P_s) \hat{Q} (A_i - B_i R_{2s}^{-1} P_s)^T] + A_P \hat{Q} + \hat{Q} A_P^T \quad (\text{A } 28)$$

$$0 = [A_P \hat{Q} + \hat{Q} (\Gamma^T A_{ca}^T G + C^T V_{2s}^{-1} Q_s^T) + Q_s V_{2s}^{-1} Q_s^T + \hat{B} V \hat{B}^T] \Gamma^T \quad (\text{A } 29)$$

$$0 = [(G^T A_{ca} \Gamma + Q_s V_{2s}^{-1} C) \hat{Q} + \hat{Q} (G^T A_{ca} \Gamma + Q_s V_{2s}^{-1} C)^T + Q_s V_{2s}^{-1} Q_s^T + \hat{B} V \hat{B}^T] \Gamma^T \quad (\text{A } 30)$$

$$0 = A_c^T P + P A_c + R_1 + L^T R L + \hat{C}^T \hat{R} \hat{C} + \sum_{i=1}^p \gamma_i \times [A_i^T P A_i + (A_i - Q_s V_{2s}^{-1} C_i)^T \hat{P} (A_i - Q_s V_{2s}^{-1} C_i)] + A_Q^T \hat{P} + \hat{P} A_Q \quad (\text{A } 31)$$

$$0 = [A_Q^T \hat{P} + \hat{P} (G^T A_{ca} \Gamma + B R_{2s}^{-1} P_s) + P_s^T R_{2s}^{-1} P_s + L^T R L + \hat{C}^T \hat{R} \hat{C}] G^T \quad (\text{A } 32)$$

$$0 = G[(G^T A_{ca} \Gamma + B R_{2s}^{-1} P_s)^T \hat{P} + \hat{P} (G^T A_{ca} \Gamma + B R_{2s}^{-1} P_s) + P_s^T R_{2s}^{-1} P_s + L^T R L + \hat{C}^T \hat{R} \hat{C}] G^T \quad (\text{A } 33)$$

Next, computing either $\Gamma(\text{A } 29) - (\text{A } 30)$ or $G(\text{A } 32) - (\text{A } 33)$ yields (6.3). Substituting this expression for A_c into (A 28), (A 29), (A 31) and (A 32) and using

$$(\text{A } 28) + G^T \Gamma(\text{A } 29)G - (\text{A } 29)G - (\text{A } 29)G^T$$

and

$$G^T \Gamma(\text{A } 29)G - (\text{A } 29)G - (\text{A } 29)G^T$$

yields (6.10) and (6.12). Using

$$(\text{A } 31) + \Gamma^T G(\text{A } 32)\Gamma - (\text{A } 32)\Gamma - (\text{A } 32)\Gamma^T$$

and

$$\Gamma^T G(\text{A } 32)\Gamma - (\text{A } 32)\Gamma - (\text{A } 32)\Gamma^T$$

yields (6.11) and (6.13).

Finally, to show that the preceding development entails no loss of generality in the optimality condition we now use (6.3)–(6.14) to obtain (A 1)–(A 8). Let $A_c, B_c, B_m, C_c, C_e, C_m, G, \Gamma, \tau, Q, P, \hat{Q}, \hat{P}, \mathbf{Q}$ be as in the statement of Theorem 6.1 and define $Q_1, Q_{12}, Q_2, P_1, P_{12}, P_2$ by (A 25)–(A 27). Using (6.2) and (6.4)–(6.8) it is easy to verify (A 4)–(A 8). Finally, substitute the definitions for $Q, P, \hat{Q}, \hat{P}, G$ and Γ into (6.10)–(6.13), reverse the steps taken earlier in the proof and use (6.3)–(6.8) to obtain (A 1) and (A 2), which completes the proof. \square

REFERENCES

- ALBERT, A., 1969, *SIAM J. Appl. Math.*, **17**, 434.
- BERNSTEIN, D. S., 1987 a, *IEEE Trans. autom. Control*, **32**, 1076-1987 b *Ibid.*, **32**, 1139.
- BERNSTEIN, D. S., and HADDAD, W. M., 1988, The Optimal Projection Equations with Petersen-Holot Bounds: Robust Controller Synthesis with Guaranteed Structured Stability Radius *IEEE Trans. autom. Control*, to be published.
- BERNSTEIN, D. S., and HYLAND, D. C., 1985, *IEEE Trans. autom. Control*, **30**, 503-1988, *J. Optim. Theory Applic.*, to be published.
- BREWER, J. W., 1978, *IEEE Trans. Circuits Syst.*, **25**, 772.
- GREELEY, S. W., and HYLAND, D. C., 1988, Reduced-Order Compensation: LQR Reduction Versus Optimal Projection *AIJAA J. Guid. Control Dynam.*, to be published.
- HADDAD, W. M., 1987, Robust optimal projection control-system synthesis. Ph.D. dissertation, Department of Mechanical Engineering, Florida Institute of Technology, Melbourne, FL.
- HADDAD, W. M., and BERNSTEIN, D. S., 1987, *IEEE Trans. autom. Control*, **32**, 1135.
- HYLAND, D. C., and BERNSTEIN, D. S., 1984, *IEEE Trans. autom. Control*, **29**, 1034-1985, *Ibid.*, **30**, 1201.
- JONCKHEERE, E. A., and SILVERMAN, L. M., 1983, *IEEE Trans. autom. Control*, **28**, 953.
- LIU, Y., and ANDERSON, B. D. O., 1986, *Int. J. Control*, **44**, 507.
- RAO, C. R., and MITRA, S. K., 1971, *Generalized Inverse of Matrices and Its Applications* (New York: Wiley).
- RICHTER, S., 1987, *Proc. American Control Conf.*, June 1987, Minneapolis, MN.
- WILSON, D. A., and KUMAR, A., 1983, *Proc. IEEE Conf. on Decision and Control*, December 1983, San Antonio, TX.
- WONHAM, W. M., 1979, *Linear Multivariable Control: a Geometric Approach* (Berlin: Springer-Verlag).

APPENDIX I: Robust Synthesis: Quadratic Bound

- 101. W. M. Haddad and D. S. Bernstein, "Robust, Reduced-Order Modeling via the Optimal Projection Equations with Petersen-Hollot Bounds," IEEE Trans. Autom. Contr., Vol. 33, pp. 692-695, 1988.
- 83. W. M. Haddad and D. S. Bernstein, "Robust, Reduced-Order, Nonstrictly Proper State Estimation via the Optimal Projection Equations with Petersen-Hollot Bounds," Sys. Contr. Lett., Vol. 9, pp. 423-431, 1987.
- 94. D. S. Bernstein and W. M. Haddad, "The Optimal Projection Equations with Petersen-Hollot Bounds: Robust Stability and Performance via Fixed-Order Dynamic Compensation for Systems with Structured Real-Valued Parameter Uncertainty," IEEE Trans. Autom. Contr., Vol. 33, pp. 578-582, 1988.
- 113. D. S. Bernstein, "Robust Stability and Performance via Fixed-Order Dynamic Compensation," SIAM J. Contr. Optim., to appear.

Robust Reduced-Order Modeling Via the Optimal Projection Equations with Petersen-Hollot Bounds

WASSIM M. HADDAD AND DENNIS S. BERNSTEIN

Abstract—An optimal reduced-order modeling problem with parametric plant uncertainty is considered. A model-reduction bound suggested by recent work of Petersen and Hollot is utilized for guaranteeing robust reduced-order modeling over a specified range of uncertain parameters. Necessary conditions which generalize the optimal projection equations for model reduction are used to characterize the reduced-order model which minimizes the model-reduction bound. The optimality equations thus effectively serve as sufficient conditions for characterizing robust reduced-order models.

I. INTRODUCTION

It has been shown in [1]–[3] that the first-order necessary conditions for quadratically optimal reduced-order modeling, estimation, and control can be transformed into coupled systems of two, three, and four matrix equations, respectively. This coupling is due to an oblique projection which arises as a direct consequence of optimality. In a series of papers [4]–[6] the optimal projection approach was generalized to the problems

of robust reduced-order estimation and control in the presence of real-valued, structured parameter uncertainty. This was accomplished by incorporating the quadratic uncertainty bound proposed in [7] within the optimal projection framework.

The purpose of the present note is to complete this cycle of results by similarly extending the results of [1]. Our goal is thus to obtain robust reduced-order models over a specified range of parametric plant uncertainty. As in [4]–[6], the main idea is to bound the effect of the uncertain parameters on the model-reduction error over the uncertainty range and then determine a reduced-order model which minimizes the model-reduction bound. The resulting generalization of the optimal projection equations now serves as a sufficient condition for robust model reduction by virtue of the fact that a bound on the model-reduction error is being minimized rather than the model-reduction error itself. These optimality conditions now comprise a coupled system of three algebraic matrix equations which reduce to the result of [1] when the uncertainty bounds are absent.

II. NOTATION AND DEFINITIONS

| | |
|---|---|
| $\mathcal{H}, \mathcal{H}^{**}, \mathcal{H}'$ | real numbers, $r \times s$ real matrices, \mathcal{H}^{**} |
| $I_r, (\cdot)^T, \mathbb{E}$ | $r \times r$ identity matrix, transpose, expected value |
| \bullet | Kronecker sum |
| $\mathcal{S}', \mathcal{P}', \mathcal{P}'$ | $r \times r$ symmetric, nonnegative-definite, positive-definite matrices |
| $Z_1 \leq Z_2, Z_1 < Z_2$ | $Z_2 - Z_1 \in \mathcal{H}', Z_2 - Z_1 \in \mathcal{P}', Z_1, Z_2 \in \mathcal{S}'$ |
| $n, l, n_m, m; \bar{n}$ | positive integers; $n + n_m$ |
| x, y, y_m, x_m, \bar{x} | n, l, l, n_m, \bar{n} -dimensional vectors |
| $A, \Delta A; B, C$ | $n \times n$ matrices, $n \times m, l \times n$ matrices |
| A_m, B_m, C_m | $n_m \times n_m, n_m \times m, l \times n_m$ matrices |
| $\bar{A}, \Delta \bar{A}, \bar{B}$ | $\begin{bmatrix} A & 0 \\ 0 & A_m \end{bmatrix}, \begin{bmatrix} \Delta A & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} B \\ B_m \end{bmatrix}$ |
| R | model-reduction error-weighting matrix in \mathcal{P}' |
| $w(\cdot)$ | m -dimensional white noise |
| V | intensity of $w(\cdot)$ in \mathcal{P}' |
| \bar{R} | $\begin{bmatrix} C^T R C & -C^T R C_m \\ -C_m^T R C & C_m^T R C_m \end{bmatrix}$ |
| \bar{P} | $\begin{bmatrix} B V B^T & B V B_m^T \\ B_m V B^T & B_m V B_m^T \end{bmatrix}$ |

III. ROBUST MODEL-REDUCTION PROBLEM

Let $\mathcal{U} \subset \mathcal{H}^{**}$ denote the set of uncertain perturbations ΔA of the nominal plant matrix A .

Robust Model-Reduction Problem: For fixed $n_m \leq n$, determine (A_m, B_m, C_m) such that, for the system consisting of the n th-order disturbed plant

$$\dot{x}(t) = (A + \Delta A)x(t) + Bw(t), \quad t \in [0, \infty), \quad (3.1)$$

outputs

$$y(t) = Cx(t), \quad (3.2)$$

and n_m th-order model

$$\dot{x}_m(t) = A_m x_m(t) + B_m w(t), \quad (3.3)$$

$$y_m(t) = C_m x_m(t), \quad (3.4)$$

the model-reduction criterion

$$J(A_m, B_m, C_m) \triangleq \sup_{\Delta A \in \mathcal{U}} \limsup_{t \rightarrow \infty} \mathbb{E}[y(t) - y_m(t)]^T R [y(t) - y_m(t)] \quad (3.5)$$

is minimized.

Manuscript received August 4, 1987; revised October 16, 1987. This work was supported in part by the Air Force Office of Scientific Research under Contract F49620-86-C-0002.

W. M. Haddad is with the Department of Mechanical Engineering, Florida Institute of Technology, Melbourne, FL 32901.

D. S. Bernstein is with the Government Aerospace Systems Division, Harris Corporation, Melbourne, FL 32902.

IEEE Log Number 8821012.

For each reduced-order model (A_m, B_m, C_m) and system variation $\Delta A \in \mathcal{U}$, the augmented system (3.1)–(3.4) is given by

$$\dot{\tilde{x}}(t) = (\tilde{A} + \Delta \tilde{A})\tilde{x}(t) + \tilde{B}w(t), \quad t \in [0, \infty) \quad (3.6)$$

where $\tilde{x}(t) \triangleq [x^T(t), x_m^T(t)]^T$.

IV. SUFFICIENT CONDITIONS FOR ROBUST PERFORMANCE

The following result is immediate.

Lemma 4.1: Suppose A_m is asymptotically stable and $A + \Delta A$ is asymptotically stable for all $\Delta A \in \mathcal{U}$. Then

$$J(A_m, B_m, C_m) = \sup_{\Delta A \in \mathcal{U}} \text{tr } \tilde{Q}_{\Delta A} \tilde{A} \quad (4.1)$$

where $\tilde{Q}_{\Delta A} \triangleq \lim_{t \rightarrow \infty} \mathbb{E}[\tilde{x}(t)\tilde{x}(t)^T] \in \mathbb{H}^d$ is the unique solution to

$$0 = (\tilde{A} + \Delta \tilde{A})\tilde{Q}_{\Delta A} + \tilde{Q}_{\Delta A}(\tilde{A} + \Delta \tilde{A})^T + \tilde{P}. \quad (4.2)$$

We now determine an upper bound for J given by (4.1).

Theorem 4.1: Let $\Omega: \mathbb{H}^d \rightarrow \mathbb{S}^d$ be such that

$$\Delta \tilde{A}Q + Q\Delta \tilde{A}^T \leq \Omega(Q), \quad \Delta A \in \mathcal{U}, Q \in \mathbb{H}^d \quad (4.3)$$

and, for given (A_m, B_m, C_m) , suppose there exists $Q \in \mathbb{H}^d$ satisfying

$$0 = \tilde{A}Q + Q\tilde{A}^T + \Omega(Q) + \tilde{P} \quad (4.4)$$

and suppose the pair $(\tilde{P}^{1/2}, \tilde{A} + \Delta \tilde{A})$ is stabilizable for all $\Delta A \in \mathcal{U}$. Then A_m is asymptotically stable, $A + \Delta A$ is asymptotically stable for all $\Delta A \in \mathcal{U}$.

$$\tilde{Q}_{\Delta A} \leq Q, \quad \Delta A \in \mathcal{U}, \quad (4.5)$$

where $\tilde{Q}_{\Delta A}$ satisfies (4.2), and

$$J(A_m, B_m, C_m) \leq \text{tr } Q\tilde{R}. \quad (4.6)$$

Proof: See [5]. \square

Remark 4.1: Theorem 4.1 provides sufficient conditions for reduced-order modeling with an upper bound on modeling error. The result yields, in addition, the result that A_m and $A + \Delta A$ are asymptotically stable. Thus, it is important to emphasize that our results are effectively limited to systems which remain asymptotically stable over the class of uncertainties. Relevant applications include, for example, damped flexible structures with uncertain modal data.

V. UNCERTAINTY STRUCTURE AND THE PETERSEN-HOLLOT BOUND

The uncertainty set \mathcal{U} is assumed to be of the form

$$\mathcal{U} = \left\{ \Delta A \in \mathbb{H}^{n \times n} : \Delta A = \sum_{i=1}^p D_i M_i N_i E_i, \quad M_i M_i^T \leq \tilde{M}_i, \quad N_i^T N_i \leq \tilde{N}_i, \right. \\ \left. i = 1, \dots, p \right\} \quad (5.1)$$

where, for $i = 1, \dots, p$: $D_i \in \mathbb{H}^{n \times n}$ and $E_i \in \mathbb{H}^{n \times n}$ are fixed matrices denoting the structure of the uncertainty; $\tilde{M}_i \in \mathbb{H}^{n \times n}$ and $\tilde{N}_i \in \mathbb{H}^{n \times n}$ are given uncertainty bounds; and $M_i \in \mathbb{H}^{n \times n}$, $N_i \in \mathbb{H}^{n \times n}$ are uncertain matrices. The augmented system thus has structured uncertainty of the form

$$\Delta \tilde{A} = \sum_{i=1}^p \tilde{D}_i M_i N_i \tilde{E}_i \quad (5.2)$$

where

$$\tilde{D}_i \triangleq \begin{bmatrix} D_i \\ 0 \end{bmatrix}, \quad \tilde{E}_i \triangleq [E_i \quad 0]. \quad (5.3)$$

We now specify the function Ω satisfying (4.3).

Proposition 5.1: The function

$$\Omega(Q) \triangleq \sum_{i=1}^p \tilde{D}_i \tilde{M}_i \tilde{D}_i^T + Q \tilde{E}_i^T \tilde{N}_i \tilde{E}_i Q \quad (5.4)$$

satisfies (4.3) with \mathcal{U} given by (5.1).

Proof: For $i = 1, \dots, p$,

$$0 \leq [\tilde{D}_i M_i - Q \tilde{E}_i^T \tilde{N}_i^T][\tilde{D}_i M_i - Q \tilde{E}_i^T \tilde{N}_i^T]^T \\ = \tilde{D}_i M_i M_i^T \tilde{D}_i^T + Q \tilde{E}_i^T \tilde{N}_i^T \tilde{N}_i \tilde{E}_i Q - (\tilde{D}_i M_i N_i \tilde{E}_i Q + Q \tilde{E}_i^T \tilde{N}_i^T M_i^T \tilde{D}_i^T) \\ \leq \tilde{D}_i \tilde{M}_i \tilde{D}_i^T + Q \tilde{E}_i^T \tilde{N}_i \tilde{E}_i Q - (\tilde{D}_i M_i N_i \tilde{E}_i Q + Q \tilde{E}_i^T \tilde{N}_i^T M_i^T \tilde{D}_i^T).$$

Summing over i yields (4.3). \square

Remark 5.1: The bound (5.4) was originally used in [7] for unit-rank perturbations with scalar uncertain parameters. For further details, see [4]–[6].

VI. THE AUXILIARY MINIMIZATION PROBLEM

Our goal is to minimize the error bound (4.6).

Auxiliary Minimization Problem: Determine (Q, A_m, B_m, C_m) with $Q \in \mathbb{H}^d$ which minimize

$$J(Q, A_m, B_m, C_m) \triangleq \text{tr } Q\tilde{R} \quad (6.1)$$

subject to

$$0 = \tilde{A}Q + Q\tilde{A}^T + \sum_{i=1}^p [\tilde{D}_i \tilde{M}_i \tilde{D}_i^T + Q \tilde{E}_i^T \tilde{N}_i \tilde{E}_i Q] + \tilde{P} \quad (6.2)$$

and

$$(\tilde{P}^{1/2}, \tilde{A} + \Delta \tilde{A}) \text{ is stabilizable, } \Delta A \in \mathcal{U}. \quad (6.3)$$

Proposition 6.1: If (Q, A_m, B_m, C_m) satisfy (6.2) and (6.3) with $Q \geq 0$, then A_m is asymptotically stable, $A + \Delta A$ is asymptotically stable for all $\Delta A \in \mathcal{U}$, and

$$J(A_m, B_m, C_m) \leq J(Q, A_m, B_m, C_m). \quad (6.4)$$

Proof: With Ω given by (5.4), Proposition 5.1 implies that (4.3) is satisfied. Hence, with (6.3), the hypotheses of Theorem 4.1 are satisfied so that the system (3.6) is stable over \mathcal{U} with model-reduction bound (4.6). Note that with (6.1), (6.4) is merely a restatement of (4.6). \square

VII. NECESSARY CONDITIONS FOR THE AUXILIARY MINIMIZATION PROBLEM

Rigorous application of the Lagrange multiplier technique requires additional technical assumptions. Specifically, we further restrict (Q, A_m, B_m, C_m) to the open set

$$\delta \triangleq \{(Q, A_m, B_m, C_m) : Q \in \mathbb{P}^d, \tilde{A} \text{ is asymptotically stable,} \\ \text{and } (A_m, B_m, C_m) \text{ is controllable and observable}\}$$

where

$$\tilde{A} \triangleq \left(\tilde{A} + \sum_{i=1}^p \tilde{E}_i^T \tilde{N}_i \tilde{E}_i Q \right) \oplus \left(\tilde{A} + \sum_{i=1}^p \tilde{E}_i^T \tilde{N}_i \tilde{E}_i Q \right).$$

Remark 7.1: The constraint $(Q, A_m, B_m, C_m) \in \delta$ is not required for robust reduced-order modeling since, as shown by Proposition 6.1, only (6.2) and (6.3) are required. As will be seen from the proof of Theorem 7.1, the set δ constitutes sufficient conditions under which the Lagrange multiplier technique is applicable to the auxiliary minimization problem. Specifically, $Q > 0$ replaces $Q \geq 0$ by an open set constraint, asymptotic stability of \tilde{A} serves as a normality condition which further implies that the dual \mathcal{P} of \mathcal{Q} is nonnegative definite, and (A_m, B_m, C_m) minimal is a nondegeneracy condition.

The following factorization lemma is needed for the statement of the main result. For details, see [1].

Lemma 7.1: If $\hat{Q}, \hat{P} \in \mathbb{R}^n$ and $\text{rank } \hat{Q}\hat{P} = n_m$, then there exist $n_m \times n$ G, Γ , and $n_m \times n_m$ invertible M such that

$$\hat{Q}\hat{P} = G^T M \Gamma, \quad (7.1)$$

$$\Gamma G^T = I_{n_m}. \quad (7.2)$$

Furthermore, G, M , and Γ are unique except for a change of basis in \mathbb{R}^{n_m} . Recall from [1] that

$$\tau \triangleq \hat{Q}\hat{P}(\hat{Q}\hat{P})^\circ = G^T \Gamma \quad (7.3)$$

is an oblique projection, where $(\cdot)^\circ$ denotes group generalized inverse. Define the complementary projection $\tau_\perp \triangleq I_n - \tau$ and call (G, M, Γ) satisfying (7.1), (7.2) a *projective factorization* of $\hat{Q}\hat{P}$. Furthermore, define the notation

$$D \triangleq \sum_{i=1}^p D_i \tilde{M}_i D_i^T, \quad E \triangleq \sum_{i=1}^p E_i^T \tilde{N}_i E_i.$$

Theorem 7.1: Suppose $(Q, A_m, B_m, C_m) \in \mathcal{S}$ solves the auxiliary minimization problem with \mathcal{U} given by (5.1). Then there exist $Q, \hat{Q}, \hat{P} \in \mathbb{R}^n$ such that Q, A_m, B_m, C_m are given by

$$Q = \begin{bmatrix} Q + \hat{Q} & \hat{Q}\Gamma^T \\ \Gamma\hat{Q} & \Gamma\hat{Q}\Gamma^T \end{bmatrix}, \quad (7.4)$$

$$A_m = \Gamma(A + QE)G^T, \quad (7.5)$$

$$B_m = \Gamma B, \quad (7.6)$$

$$C_m = CG^T \quad (7.7)$$

for some projective factorization (G, M, Γ) of $\hat{Q}\hat{P}$, and such that Q, \hat{Q}, \hat{P} satisfy

$$0 = A Q + Q A^T + D + Q E Q + \tau_\perp B V B^T \tau_\perp^T, \quad (7.8)$$

$$0 = (A + QE)\hat{Q} + \hat{Q}(A + QE)^T + \hat{Q}E\hat{Q} + B V B^T - \tau_\perp B V B^T \tau_\perp^T, \quad (7.9)$$

$$0 = (A + QE)^T \hat{P} + \hat{P}(A + QE) + C^T R C - \tau_\perp^T C^T R C \tau_\perp, \quad (7.10)$$

$$\text{rank } \hat{Q} = \text{rank } \hat{P} = \text{rank } \hat{Q}\hat{P} = n_m. \quad (7.11)$$

Furthermore, the auxiliary cost is given by

$$\mathcal{J}(Q, A_m, B_m, C_m) = \text{tr } Q C^T R C. \quad (7.12)$$

Conversely, if $Q, \hat{Q}, \hat{P} \in \mathbb{R}^n$ satisfy (7.8)–(7.11) then (Q, A_m, B_m, C_m) given by (7.4)–(7.7) satisfy $Q \geq 0$ and (6.2) with auxiliary cost given by (7.12).

Proof: See the Appendix. \square

Theorem 7.1 presents necessary conditions for the auxiliary minimization problem which explicitly characterize extremals (Q, A_m, B_m, C_m) . These necessary conditions consist of a system of two modified Lyapunov equations and one modified Riccati equation coupled by an oblique projection τ and uncertainty terms. Setting D and E to zero, i.e., deleting the plant uncertainties, it can be seen that (7.8) drops out while (7.9) and (7.10) reduce to the optimal projection equations for model reduction obtained in [1]. If, alternatively, $n_m = n$, then the full-order robust model is given by $A + QE, B, C$ where Q is given by (7.8) with $\tau_\perp = 0$.

Remark 7.2: As in the perfect modeling case considered in [1], (7.8)–(7.10) may support multiple solutions. When uncertainty is present but a full-order model is desired, then the solution is unique.

Remark 7.3: The conservatism of the bound (7.12) is difficult to predict for two reasons. First, the overbounding (4.3) holds with respect to the partial ordering of the nonnegative-definite matrices for which no scalar measure of conservatism is available. And, second, the bound (4.3) is required to hold for all nonnegative-definite matrices Q . The conservatism will thus depend upon the actual value of Q determined by solving (6.2). Numerical experience with related bounds shows that the conservatism is highly problem dependent. See [8].

VIII. SUFFICIENT CONDITIONS FOR ROBUST REDUCED-ORDER MODELING

The main result guaranteeing robust model reduction can now be stated.

Theorem 8.1: Suppose there exist $Q, \hat{Q}, \hat{P} \in \mathbb{R}^n$ satisfying (7.8)–(7.11) and suppose that $(P^{1/2}, \hat{A} + \Delta\hat{A})$ is stabilizable for all $\Delta\hat{A} \in \mathcal{U}$ with A_m, B_m, C_m given by (7.5)–(7.7) and \mathcal{U} defined by (5.1). Then A_m is asymptotically stable, $A + \Delta A$ is asymptotically stable for all $\Delta A \in \mathcal{U}$, and the model-reduction criterion satisfies the bound

$$\mathcal{J}(A_m, B_m, C_m) \leq \text{tr } Q C^T R C. \quad (8.1)$$

Proof: The converse of Theorem 7.1 implies that Q given by (7.4) is nonnegative definite and satisfies (6.2). With the stabilizability assumption the result follows from Proposition 6.1. \square

Remark 8.1: As noted in Remark 4.1, Theorem 8.1 is effectively limited to systems which remain asymptotically stable over the class of uncertainties.

APPENDIX

PROOF OF THEOREM 7.1

To optimize (6.1) over the open set

$$\mathcal{S}' \triangleq \{(Q, A_m, B_m, C_m) \in \mathcal{S} : (6.3) \text{ is satisfied}\}$$

subject to the constraint (6.2), form the Lagrangian

$$\mathcal{L}(A_m, B_m, C_m, Q, \Phi, \lambda) \triangleq$$

$$\text{tr} \left\{ \lambda Q R + \left[\lambda Q + Q \lambda^T + \sum_{i=1}^p D_i \tilde{M}_i D_i^T + Q E^T \tilde{N}_i E_i Q + V \right] \Phi \right\},$$

where the Lagrange multipliers $\lambda \geq 0$ and $\Phi \in \mathbb{R}^{n \times n}$ are not both zero. We thus obtain

$$\frac{\partial \mathcal{L}}{\partial Q} = \tilde{A}^T \Phi + \Phi \tilde{A} + \sum_{i=1}^p [E_i^T \tilde{N}_i E_i Q \Phi + \Phi Q E_i^T \tilde{N}_i E_i] + \lambda R.$$

Setting $\partial \mathcal{L} / \partial Q = 0$ yields

$$\tilde{Q}^T \text{vec } \Phi = -\lambda \text{vec } \tilde{R} \quad (A.1)$$

where "vec" is the column-stacking operation (see [6]). Since \tilde{Q} is assumed to be stable, \tilde{Q}^T is invertible, and thus $\lambda = 0$ implies $\Phi = 0$. Hence, it can be assumed without loss of generality that $\lambda = 1$. Furthermore, the stability of \tilde{Q}^T implies that Φ is nonnegative definite.

Now partition $\tilde{n} \times \tilde{n}$ Q, Φ into $n \times n$, $n \times n_m$, and $n_m \times n_m$ subblocks as

$$Q = \begin{bmatrix} Q_1 & Q_{12} \\ Q_{12}^T & Q_2 \end{bmatrix}, \quad \Phi = \begin{bmatrix} P_1 & P_{12} \\ P_{12}^T & P_2 \end{bmatrix}.$$

Thus, the stationarity conditions are given by

$$\frac{\partial \mathcal{L}}{\partial Q} = \tilde{A}^T \Phi + \Phi \tilde{A} + \sum_{i=1}^p [E_i^T \tilde{N}_i E_i Q \Phi + \Phi Q E_i^T \tilde{N}_i E_i] + \tilde{R} = 0, \quad (A.2)$$

$$\frac{\partial \mathcal{L}}{\partial A_m} = P_{12}^T Q_{12} + P_2 Q_2 = 0, \quad (A.3)$$

$$\frac{\partial \mathcal{L}}{\partial B_m} = P_{12}^T B V + P_2 B_m V = 0, \quad (A.4)$$

$$\frac{\partial \mathcal{L}}{\partial C_m} = -R C_{12} + R C_m Q_2 = 0. \quad (A.5)$$

Expanding (6.2) and (A.2) yields

$$0 = AQ_1 + Q_1 A^T + D + Q_1 EQ_1 + BVB^T, \quad (A.6)$$

$$0 = AQ_{12} + Q_{12} A_m^T + Q_1 EQ_{12} + BVB_m^T, \quad (A.7)$$

$$0 = A_m Q_2 + Q_2 A_m^T + Q_{12}^T EQ_{12} + B_m VB_m^T, \quad (A.8)$$

$$0 = A^T P_1 + P_1 A + E(P_1 Q_1 + P_{12} Q_{12})^T + (P_1 Q_1 + P_{12} Q_{12})E + C^T RC, \quad (A.9)$$

$$0 = A^T P_{12} + P_{12} A_m + E(P_{12}^T Q_1 + P_2 Q_{12}^T)^T - C^T RC_m, \quad (A.10)$$

$$0 = A_m^T P_2 + P_2 A_m + C_m^T RC_m. \quad (A.11)$$

Now define the $n \times n$ nonnegative-definite matrices

$$Q = Q_1 - Q_{12} Q_2^{-1} Q_{12}^T, \quad P = P_1 - P_{12} P_2^{-1} P_{12}^T,$$

$$\hat{Q} = Q_{12} Q_2^{-1} Q_{12}^T, \quad \hat{P} = P_{12} P_2^{-1} P_{12}^T$$

and the $n_m \times n$, $n_m \times n_m$, $n_m \times n$ matrices

$$G = Q_2^{-1} Q_{12}^T, \quad M = Q_2 P_2, \quad \Gamma = -P_2^{-1} P_{12}^T.$$

The existence of Q_2^{-1} and P_2^{-1} follows from the fact that (A_m, B_m, C_m) is minimal. See [1]–[4], [6] for details.

Note that (A.3) implies (7.1) and (7.2). Sylvester's inequality yields (7.11). Next (7.4), (7.6), and (7.7) follow from the definition of Q , relations (A.4) and (A.5), and the identities

$$Q_1 = Q + \hat{Q}, \quad Q_{12} = \hat{Q} \Gamma^T, \quad P_{12} = -\hat{P} \Gamma^T, \quad Q_2 = \Gamma \hat{Q} \Gamma^T, \quad P_2 = G \hat{P} G^T.$$

Computing either $\Gamma(A.7)$, (A.8) or $G(A.10) + (A.11)$ yields (7.5). Inserting (7.5)–(7.7) into (A.6)–(A.11) it can be shown that (A.8), (A.9), and (A.11) are superfluous. Using (A.6) + $G^T \Gamma(A.7)G - (A.7)G - [(A.7)G]^T$ and $G^T \Gamma(A.7)G - (A.7)G - [(A.7)G]^T$ yield (7.8) and (7.9). Similarly, $\Gamma^T G(A.10)\Gamma - (A.10)\Gamma - [(A.10)\Gamma]^T$ yields (7.10).

Finally, the proof can be reversed so that (7.5)–(7.11) yield (A.1)–(A.5) and (6.2).

REFERENCES

- [1] D. C. Hyland and D. S. Bernstein, "The optimal projection equations for model reduction and the relationships among the methods of Wilson, Skelton, and Moore," *IEEE Trans. Automat. Contr.*, vol. AC-30, pp. 1201–1211, 1985.
- [2] D. S. Bernstein and D. C. Hyland, "The optimal projection equations for reduced-order state estimation," *IEEE Trans. Automat. Contr.*, vol. AC-30, pp. 583–585, 1985.
- [3] D. C. Hyland and D. S. Bernstein, "The optimal projection equations for fixed-order dynamic compensation," *IEEE Trans. Automat. Contr.*, vol. AC-29, pp. 1034–1037, 1984.
- [4] W. M. Haddad and D. S. Bernstein, "Robust, reduced-order, nonstrictly proper state estimation via the optimal projection equations with Petersen-Hollot bounds," *Syst. Contr. Lett.*, vol. 9, pp. 423–431, 1987.
- [5] D. S. Bernstein and W. M. Haddad, "The optimal projection equations with Petersen-Hollot Bounds: Robust stability and performance via fixed-order dynamic compensation for systems with structured real-valued parameter variations," *IEEE Trans. Automat. Contr.*, vol. 33, pp. 578–582, 1988.
- [6] D. S. Bernstein and W. M. Haddad, "The optimal projection equations with Petersen-Hollot bounds: Robust controller synthesis with guaranteed structured stability radius," in *Proc. IEEE Conf. Decision Contr.*, Los Angeles, CA, Dec. 1987, pp. 1308–1318.
- [7] I. R. Petersen and C. V. Hollot, "A Riccati equation approach to the stabilization of uncertain systems," *Automatica*, vol. 22, pp. 397–411, 1986.
- [8] D. S. Bernstein and W. M. Haddad, "Robust stability and performance analysis for linear dynamic systems," submitted for publication.

Defect Correction Methods for the Solution of Algebraic Riccati Equations

V. MEHRMANN AND E. TAN

Abstract—The solution of discrete and continuous algebraic Riccati equations is considered. It is shown that if an approximate solution is obtained, then the defect for this solution again solves an algebraic Riccati equation of the same form and that the system properties of detectability and stabilizability are inherited by this defect equation. On the basis of these results, a general defect correction method is proposed and numerical examples are given for the use of this method in combination with the SR method.

I. INTRODUCTION

We consider the numerical solution of generalized algebraic Riccati equations

$$0 = A^* X E + E^* X A - (B^* X E + S^*) R^{-1} (B^* X E + S^*) + C^* Q C = \mathcal{C}R(X) \quad (1.1)$$

$$0 = -E^* X E + A^* X A - (A^* X B + S) (R + B^* X B)^{-1} \cdot (A^* X B + S)^* + C^* Q C = \mathcal{D}R(X) \quad (1.2)$$

where $X, A, E \in \mathbb{C}^{n,n}$, $B, S \in \mathbb{C}^{n,m}$, $C \in \mathbb{C}^{r,n}$, $Q = Q^* \in \mathbb{C}^{r,r}$, $R = R^* \in \mathbb{C}^{m,m}$ are positive definite and E is nonsingular. Both equations arise, for example, in the solutions of linear quadratic optimal control problems. Equation (1.1) stems from a continuous-time and (1.2) from a discrete-time problem (see, e.g., [1], [4], and [5]). The numerical solution of these two equations has been studied extensively in recent years (e.g., [7], [4]–[6], [9], [14], [16], [17], and [20]). Typically, solutions are obtained using QR- or QZ-type algorithms to compute deflating subspaces of the matrix pencils

$$\begin{bmatrix} A & 0 & B \\ C^* Q C & A^* & S \\ S^* & B^* & R \end{bmatrix} - \lambda \begin{bmatrix} E & 0 & 0 \\ 0 & -E^* & 0 \\ 0 & 0 & 0 \end{bmatrix} \quad (1.3)$$

corresponding to problem (1.1) and

$$\begin{bmatrix} A & 0 & B \\ C^* Q C & -E^* & S \\ S^* & 0 & R \end{bmatrix} - \lambda \begin{bmatrix} E & 0 & 0 \\ 0 & -A^* & 0 \\ 0 & -B^* & 0 \end{bmatrix} \quad (1.4)$$

corresponding to problem (1.2), where, using the fact that R is positive definite, many of the algorithms are applied to the reduced pencils

$$\begin{bmatrix} F & G \\ H & -F^* \end{bmatrix} - \lambda \begin{bmatrix} E & 0 \\ 0 & E^* \end{bmatrix} = \begin{bmatrix} A - BR^{-1}S^* & BR^{-1}B^* \\ C^* Q C - SR^{-1}S^* & -A^* + SR^{-1}B^* \end{bmatrix} - \lambda \begin{bmatrix} E & 0 \\ 0 & E^* \end{bmatrix} \quad (1.5)$$

corresponding to (1.3) and

$$\begin{bmatrix} F & 0 \\ H & E^* \end{bmatrix} - \lambda \begin{bmatrix} E & -G \\ 0 & F^* \end{bmatrix} = \begin{bmatrix} A - BR^{-1}S & 0 \\ C^* Q C - SR^{-1}S^* & E^* \end{bmatrix} - \lambda \begin{bmatrix} E & -BR^{-1}B^* \\ 0 & A^* - SR^{-1}B^* \end{bmatrix} \quad (1.6)$$

corresponding to (1.4).

Manuscript received February 13, 1987; revised September 28, 1987. Paper recommended by Associate Editor, B. H. Krogh.

The authors are with the Fakultät für Mathematik, Universität Bielefeld, Bielefeld, West Germany.

IEEE Log Number 8821013.

Robust, reduced-order, nonstrictly proper state estimation via the optimal projection equations with Petersen-Hollot bounds

Wassim M. HADDAD

Department of Mechanical Engineering, Florida Institute of Technology, Melbourne, FL 32901, USA

Dennis S. BERNSTEIN *

Harris Corporation, Government Aerospace Systems Division, MS 22/4848, Melbourne, FL 32902, USA

Received 10 March 1987

Revised 11 June 1987

Abstract: A state-estimation design problem involving parametric plant uncertainties is considered. An error bound suggested by recent work of Petersen and Hollot is utilized for guaranteeing robust estimation. Necessary conditions which generalize the optimal projection equations for reduced-order state estimation are used to characterize the estimator which minimizes the error bound. The design equations thus effectively serve as sufficient conditions for synthesizing robust estimators. An additional feature is the presence of a static estimation gain in conjunction with the dynamic (Kalman) estimator, i.e., a nonstrictly proper estimator.

Keywords: Robust Kalman filter, Error bounds, Reduced-order state estimation.

1. Introduction

As is well known [2,5-8,11,14,15] optimal filters based upon nominal parameter values may be severely degraded in the presence of parameter deviations. Thus, it is desirable to obtain robust state estimators which provide acceptable performance over the range of parametric uncertainty. The approach of the present paper is related to the guaranteed cost approach developed for control in [4,16] and applied to estimation in [11]. Specifically, the main idea is to bound the effect of the uncertain parameters on the estimation error over the uncertainty range and then choose estimator gains to minimize the estimation bound. Thus the actual estimation error is guaranteed to lie below the prescribed upper bound.

The technique used to determine minimizing estimator gains is based upon a generalization of the optimal projection equations for reduced-order state estimation [1]. Thus the results of the present paper effectively extend the results of [1] to the case of system uncertainties. It should be noted that the optimal projection equations, which are necessary conditions for optimality, now serve as sufficient conditions for robust estimation by virtue of the fact that a bound on the estimation error is being minimized rather than the estimation error itself. The bound utilized in the present paper is an extension of the approach developed in [12,13] for constructing Lyapunov functions for full-state feedback and utilized in [10] to characterize the structured stability radius.

An additional feature of the present paper is the inclusion of a static feedback gain in conjunction with the dynamic estimator. Thus the results of the present paper represent a generalization of standard results to the case of nonstrictly proper estimation.

* Supported in part by the Air Force Office of Scientific Research under contract F49620-86-C-0002.

2. Notation and definitions

Note: All matrices have real entries.

| | |
|---|--|
| $\mathbf{R}, \mathbf{R}^{r \times s}, \mathbf{R}^r, \mathbf{E}$ | real numbers, $r \times s$ real matrices, $\mathbf{R}^{r \times 1}$, expected value. |
| $I_r, (\cdot)^T$ | $r \times r$ identity matrix, transpose |
| $\mathbf{S}^r, \mathbf{N}^r, \mathbf{P}^r$ | $r \times r$ symmetric, nonnegative-definite, positive-definite matrices. |
| $Z_1 \leq Z_2, Z_1 < Z_2$ | $Z_2 - Z_1 \in \mathbf{N}^r, Z_2 - Z_1 \in \mathbf{P}^r, Z_1, Z_2 \in \mathbf{S}^r$. |
| $n, l, \hat{l}, n_e, p, q; \tilde{n}$ | positive integers; $n + n_e$. |
| $x, y, \hat{y}, y_e, x_e, \tilde{x}$ | $n, l, \hat{l}, q, n_e, \tilde{n}$ -dimensional vectors. |
| $A, \Delta A; C, \Delta C$ | $n \times n$ matrices; $l \times n$ matrices. |
| \hat{C}, L, R | $\hat{l} \times n$ matrix, $q \times n$ matrix, $q \times q$ matrix, $R \in \mathbf{P}^q$. |
| A_e, B_e, C_e, D_e | $n_e \times n_e, n_e \times l, q \times n_e, q \times l$ matrices. |
| $\tilde{A}, \Delta \tilde{A}$ | $\begin{bmatrix} A & 0 \\ B_e C & A_e \end{bmatrix}, \begin{bmatrix} \Delta A & 0 \\ B_e \Delta C & 0 \end{bmatrix}$. |
| \tilde{R} | $\begin{bmatrix} L^T R L - L^T R D_e \hat{C} - \hat{C}^T D_e^T R L + \hat{C}^T D_e^T R D_e \hat{C} & -L^T R C_e + \hat{C}^T D_e^T R C_e \\ -C_e^T R L + C_e^T R D_e \hat{C} & C_e^T R C_e \end{bmatrix}$. |
| $w_0(\cdot), w_1(\cdot)$ | n, l -dimensional white noise. |
| V_0, V_1 | intensity of $w_0(\cdot), w_1(\cdot)$; $V_0 \in \mathbf{N}^n, V_1 \in \mathbf{P}^l$. |
| V_{01} | $n \times l$ cross intensity of $w_0(\cdot), w_1(\cdot)$. |
| $\tilde{w}(\cdot), \tilde{V}$ | $\begin{bmatrix} w_0(\cdot) \\ B_e w_1(\cdot) \end{bmatrix}, \begin{bmatrix} V_0 & V_{01} B_e^T \\ B_e V_{01}^T & B_e V_1 B_e^T \end{bmatrix}$. |

3. Robust estimation problem

Let $\mathcal{U} \subset \mathbb{R}^{n \times n} \times \mathbb{R}^{l \times n}$ denote the set of uncertain perturbations $(\Delta A, \Delta C)$ of the nominal plant matrices A and C .

Robust Estimation Problem. For fixed $n_e \leq n$, determine (A_e, B_e, C_e, D_e) such that, for the system consisting of the n -th-order disturbed plant

$$\dot{x}(t) = (A + \Delta A)x(t) + w_0(t), \quad t \in [0, \infty), \quad (3.1)$$

noisy and nonnoisy measurements

$$y(t) = (C + \Delta C)x(t) + w_1(t), \quad (3.2)$$

$$\hat{y}(t) = \hat{C}x(t), \quad (3.3)$$

and n_e -th-order nonstrictly proper state estimator

$$\dot{x}_e(t) = A_e x_e(t) + B_e y(t), \quad (3.4)$$

$$y_e(t) = C_e x_e(t) + D_e \hat{y}(t), \quad (3.5)$$

the state-estimation error criterion

$$J(A_e, B_e, C_e, D_e) \triangleq \sup_{(\Delta A, \Delta C) \in \mathcal{U}} \limsup_{t \rightarrow \infty} \mathbb{E} [Lx(t) - y_e(t)]^T R [Lx(t) - y_e(t)] \quad (3.6)$$

is minimized.

Note that the augmented system (3.1)–(3.5) can be written as

$$\dot{\tilde{x}}(t) = (\tilde{A} + \Delta \tilde{A})\tilde{x}(t) + \tilde{w}(t), \quad t \in [0, \infty), \quad (3.7)$$

where $\tilde{x}(t) \triangleq [x^T(t), x_e^T(t)]^T$. The cost can be expressed in terms of the augmented second-moment matrix.

Proposition 3.1. For given (A_e, B_e, C_e, D_e) and $(\Delta A, \Delta C) \in \mathcal{U}$ the second-moment matrix

$$\tilde{Q}_{\Delta \tilde{A}}(t) \triangleq \mathbb{E}[\tilde{x}(t)\tilde{x}^T(t)], \quad t \in [0, \infty), \quad (3.8)$$

satisfies

$$\dot{\tilde{Q}}_{\Delta \tilde{A}}(t) = (\tilde{A} + \Delta \tilde{A})\tilde{Q}_{\Delta \tilde{A}}(t) + \tilde{Q}_{\Delta \tilde{A}}(t)(\tilde{A} + \Delta \tilde{A})^T + \tilde{V}, \quad t \in [0, \infty). \quad (3.9)$$

Furthermore,

$$J(A_e, B_e, C_e, D_e) = \sup_{(\Delta A, \Delta C) \in \mathcal{U}} \limsup_{t \rightarrow \infty} \text{tr } \tilde{Q}_{\Delta \tilde{A}}(t) \tilde{R}. \quad (3.10)$$

4. Sufficient conditions for robust performance

Lemma 4.1. Suppose the system (3.7) is stable for all $(\Delta A, \Delta C) \in \mathcal{U}$. Then

$$J(A_e, B_e, C_e, D_e) = \sup_{(\Delta A, \Delta C) \in \mathcal{U}} \text{tr } \tilde{Q}_{\Delta \tilde{A}} \tilde{R}. \quad (4.1)$$

where $\tilde{Q}_{\Delta \tilde{A}} \in \mathbb{N}^{\tilde{n}}$ is the unique solution to

$$0 = (\tilde{A} + \Delta \tilde{A})\tilde{Q}_{\Delta \tilde{A}} + \tilde{Q}_{\Delta \tilde{A}}(\tilde{A} + \Delta \tilde{A})^T + \tilde{V}. \quad (4.2)$$

We now seek upper bounds for $J(A_e, B_e, C_e, D_e)$.

Theorem 4.1. Let $\Omega: \mathbb{N}^{\tilde{n}} \times \mathbb{R}^{n_e \times l} \rightarrow \mathbb{S}^{\tilde{n}}$ be such that

$$\Delta \tilde{A} \mathcal{Q} + \mathcal{Q} \Delta \tilde{A}^T \leq \Omega(\mathcal{Q}, B_e), \quad (\Delta A, \Delta C) \in \mathcal{U}, \quad (\mathcal{Q}, B_e) \in \mathbb{N}^{\tilde{n}} \times \mathbb{R}^{n_e \times l}, \quad (4.3)$$

and, for given (A_e, B_e, C_e, D_e) , suppose there exists $\mathcal{Q} \in \mathbb{N}^{\tilde{n}}$ satisfying

$$0 = \tilde{A} \mathcal{Q} + \mathcal{Q} \tilde{A}^T + \Omega(\mathcal{Q}, B_e) + \tilde{V}, \quad (4.4)$$

and suppose the pair $(\tilde{V}^{1/2}, \tilde{A} + \Delta \tilde{A})$ is detectable for all $(\Delta A, \Delta C) \in \mathcal{U}$. Then A_e is asymptotically stable, $A + \Delta A$ is asymptotically stable for all $(\Delta A, \Delta C) \in \mathcal{U}$,

$$\tilde{Q}_{\Delta \tilde{A}} \leq \mathcal{Q}, \quad (\Delta A, \Delta C) \in \mathcal{U}, \quad (4.5)$$

where $\tilde{Q}_{\Delta \tilde{A}}$ satisfies (4.2), and

$$J(A_e, B_e, C_e, D_e) \leq \text{tr } \mathcal{Q} \tilde{R}. \quad (4.6)$$

Proof. For all $(\Delta A, \Delta C) \in \mathcal{U}$, (4.4) is equivalent to

$$0 = (\tilde{A} + \Delta \tilde{A})\mathcal{Q} + \mathcal{Q}(\tilde{A} + \Delta \tilde{A})^T + \Psi(\mathcal{Q}, B_e, \Delta \tilde{A}) + \tilde{V}, \quad (4.7)$$

where

$$\Psi(\mathcal{Q}, B_e, \Delta \tilde{A}) \triangleq \Omega(\mathcal{Q}, B_e) - (\tilde{A} \mathcal{Q} + \mathcal{Q} \tilde{A}^T).$$

Note that by (4.3), $\Psi(\mathcal{Q}, B_e, \Delta \tilde{A}) \geq 0$ for all $(\Delta A, \Delta C) \in \mathcal{U}$. Since $(\tilde{V}^{1/2}, \tilde{A} + \Delta \tilde{A})$ is detectable for all $(\Delta A, \Delta C) \in \mathcal{U}$, it follows from Theorem 3.6 of [17] that $((\tilde{V} + \Psi(\mathcal{Q}, B_e, \Delta \tilde{A}))^{1/2}, \tilde{A} + \Delta \tilde{A})$ is detectable for all $(\Delta A, \Delta C) \in \mathcal{U}$. Hence Lemma 12.2 of [17] implies $\tilde{A} + \Delta \tilde{A}$ is asymptotically stable for all $(\Delta A, \Delta C) \in \mathcal{U}$. Since $\tilde{A} + \Delta \tilde{A}$ is lower block triangular, A_e is asymptotically stable and $A + \Delta A$ is asymptotically stable for all $(\Delta A, \Delta C) \in \mathcal{U}$. Next, subtracting (4.2) from (4.7) yields

$$0 = (\tilde{A} + \Delta \tilde{A})(\mathcal{Q} - \tilde{Q}_{\Delta \tilde{A}}) + (\mathcal{Q} - \tilde{Q}_{\Delta \tilde{A}})(\tilde{A} + \Delta \tilde{A})^T + \Psi(\mathcal{Q}, B_e, \Delta \tilde{A}).$$

or, equivalently, (since $\tilde{A} + \Delta\tilde{A}$ is asymptotically stable)

$$\mathcal{Q} - \tilde{\mathcal{Q}}_{\Delta\tilde{A}} = \int_0^\infty e^{(\tilde{A} + \Delta\tilde{A})t} \Psi(\mathcal{Q}, B_c, \Delta\tilde{A}) e^{(\tilde{A} + \Delta\tilde{A})^T t} dt \geq 0,$$

which implies (4.5). Finally, (4.5) and (4.1) yield (4.6). \square

5. Uncertainty structure

The uncertainty set \mathcal{U} is assumed to be of the form

$$\mathcal{U} = \left\{ (\Delta A, \Delta C) \in \mathbb{R}^{n \times n} \times \mathbb{R}^{l \times n} : \Delta A = \sum_{i=1}^p D_i M_i N_i E_i, \right. \\ \left. \Delta C = \sum_{i=1}^p F_i M_i N_i E_i, M_i M_i^T \leq \bar{M}_i, N_i^T N_i \leq \bar{N}_i, i = 1, \dots, p \right\}, \quad (5.1)$$

where, for $i = 1, \dots, p$, $D_i \in \mathbb{R}^{n \times r_i}$, $E_i \in \mathbb{R}^{l_i \times n}$ and $F_i \in \mathbb{R}^{l_i \times r_i}$ are fixed matrices denoting the structure of the uncertainty; $\bar{M}_i \in \mathbb{N}^{r_i}$ and $\bar{N}_i \in \mathbb{N}^{l_i}$ are given uncertainty bounds; and $M_i \in \mathbb{R}^{r_i \times s_i}$, $N_i \in \mathbb{R}^{s_i \times l_i}$ are uncertain matrices. The closed-loop system thus has structured uncertainty of the form

$$\Delta\tilde{A} = \sum_{i=1}^p \tilde{D}_i M_i N_i \tilde{E}_i, \quad (5.2)$$

where

$$\tilde{D}_i \triangleq \begin{bmatrix} D_i \\ B_c F_i \end{bmatrix}, \quad \tilde{E}_i \triangleq [E_i \ 0]. \quad (5.3)$$

The special case $\bar{M}_i = \mu_i^2 I_{r_i}$, $\bar{N}_i = \nu_i^2 I_{l_i}$ is worth noting.

Proposition 5.1. Let $\mu_i, \nu_i \geq 0$, $i = 1, \dots, p$. Then $M_i M_i^T \leq \mu_i^2 I_{r_i}$ and $N_i^T N_i \leq \nu_i^2 I_{l_i}$ if and only if $\sigma_{\max}(M_i) \leq \mu_i$ and $\sigma_{\max}(N_i) \leq \nu_i$.

Remark 5.1. \mathcal{U} given by (5.1) is directly related to the structured stability radius introduced in [10]. Setting $p = 1$, $\bar{M}_1 = \mu_1 I_{r_1}$, $r_1 = s_1$, $\bar{N}_1 = I_{r_1}$ and $N_1 \equiv I_{r_1}$ yields the setting of [10]. For a similar formulation, see [13].

6. The Petersen-Hollot bound

Given \mathcal{U} as defined in (5.1), we now specify Ω satisfying (4.3).

Proposition 6.1. The bound Ω given by

$$\Omega(\mathcal{Q}, B_c) \triangleq \sum_{i=1}^p \tilde{D}_i \bar{M}_i \tilde{D}_i^T + \mathcal{Q} \tilde{E}_i^T \bar{N}_i \tilde{E}_i \mathcal{Q} \quad (6.1)$$

satisfies (4.3) with \mathcal{U} given by (5.1).

Proof. For $i = 1, \dots, p$,

$$\begin{aligned} 0 &\leq [\bar{D}_i M_i - \mathcal{Q} \bar{E}_i^T N_i^T] [\bar{D}_i M_i - \mathcal{Q} \bar{E}_i^T N_i^T]^T \\ &= \bar{D}_i M_i M_i^T \bar{D}_i^T + \mathcal{Q} \bar{E}_i^T N_i^T N_i \bar{E}_i \mathcal{Q} - (\bar{D}_i M_i N_i \bar{E}_i \mathcal{Q} + \mathcal{Q} \bar{E}_i^T N_i^T M_i^T \bar{D}_i^T) \\ &\leq \bar{D}_i \bar{M}_i \bar{D}_i^T + \mathcal{Q} \bar{E}_i^T \bar{N}_i \bar{E}_i \mathcal{Q} - (\bar{D}_i M_i N_i \bar{E}_i \mathcal{Q} + \mathcal{Q} \bar{E}_i^T N_i^T M_i^T \bar{D}_i^T). \end{aligned}$$

Summing over i yields (4.3). \square

Remark 6.1. The bound (6.1) is used in [12] for unit-rank perturbations while a more general treatment appears in [13].

7. The auxiliary minimization problem

Our goal is to minimize the error bound (4.6).

Auxiliary Minimization Problem. Determine $(\mathcal{Q}, A_e, B_e, C_e, D_e)$ with $\mathcal{Q} \in \mathbb{N}^n$ which minimizes

$$\mathcal{J}(\mathcal{Q}, A_e, B_e, C_e, D_e) \triangleq \text{tr } \mathcal{Q} \bar{R} \quad (7.1)$$

subject to

$$0 = \bar{A} \mathcal{Q} + \mathcal{Q} \bar{A}^T + \sum_{i=1}^p [\bar{D}_i \bar{M}_i \bar{D}_i^T + \mathcal{Q} \bar{E}_i^T \bar{N}_i \bar{E}_i \mathcal{Q}] + \bar{V}. \quad (7.2)$$

and

$$(\bar{V}^{1/2}, \bar{A} + \Delta \bar{A}) \text{ is detectable, } (\Delta A, \Delta C) \in \mathcal{U}. \quad (7.3)$$

Proposition 7.1. If $(\mathcal{Q}, A_e, B_e, C_e, D_e)$ satisfies (7.2) and (7.3) with $\mathcal{Q} \geq 0$, then $\bar{A} + \Delta \bar{A}$ is asymptotically stable for all $(\Delta A, \Delta C) \in \mathcal{U}$, and

$$J(A_e, B_e, C_e, D_e) \leq \mathcal{J}(\mathcal{Q}, A_e, B_e, C_e, D_e). \quad (7.4)$$

Proof. With \mathcal{Q} given by (6.1), (7.2) is equivalent to (4.4). Hence, with (7.3), the hypotheses of Theorem 4.1 are satisfied so that the augmented system is stable over \mathcal{U} with estimation bound (4.6). Note that with (7.1), (7.4) is merely a restatement of (4.6). \square

8. Necessary conditions for the auxiliary minimization problem

Rigorous application of the Lagrange multiplier technique requires additional technical assumptions. Specifically, we further restrict $(\mathcal{Q}, A_e, B_e, C_e, D_e)$ to the set

$$\begin{aligned} \mathcal{S} \triangleq \{(\mathcal{Q}, A_e, B_e, C_e, D_e) : \mathcal{Q} \in \mathbb{P}^n, \bar{\mathcal{A}} \text{ is asymptotically stable, } (A_e, B_e, C_e) \text{ is controllable} \\ \text{and observable, and } \hat{C}(Q_1 - Q_{12}Q_2^{-1}Q_{12}^T)\hat{C}^T > 0\}, \end{aligned}$$

where (\oplus) denotes Kronecker sum [3])

$$\bar{\mathcal{A}} \triangleq \left(\bar{A} + \sum_{i=1}^p \bar{E}_i^T \bar{N}_i \bar{E}_i \mathcal{Q} \right) \oplus \left(\bar{A} + \sum_{i=1}^p \bar{E}_i^T \bar{N}_i \bar{E}_i \mathcal{Q} \right)$$

and \mathcal{Q} is partitioned as in Appendix A. As shown in Appendix A, Q_2 is invertible since (A_e, B_e) is controllable. The positive definiteness condition holds when \hat{C} has full row rank and \mathcal{Q} is positive definite.

As can be seen from the proof of Theorem 8.1 in Appendix A, this condition implies the existence of the projection τ_1 defined below. Note that \mathcal{S} is open.

Remark 8.1. The constraint $(\mathcal{Q}, A_e, B_e, C_e, D_e) \in \mathcal{S}$ is not required for robust estimation. As will be seen from the proof of Theorem 8.1, the set \mathcal{S} constitutes sufficient conditions under which the Lagrange multiplier technique is applicable to the Auxiliary Minimization Problem. Specifically, asymptotic stability of \mathcal{A} serves as a normality condition which further implies that the dual \mathcal{P} of \mathcal{Q} satisfying (A.2) is nonnegative definite. Furthermore, (A_e, B_e, C_e) minimal is a nondegeneracy condition which implies that the lower right $n_e \times n_e$ subblocks of \mathcal{Q} and \mathcal{P} are positive definite. It is extremely important to emphasize that Proposition 7.1 shows that it is *not* necessary for guaranteed robust estimation that an admissible quadruple obtained by solving the necessary conditions actually be shown to be an element of \mathcal{S} .

For arbitrary $Q \in \mathbb{R}^{n \times n}$ define the following notation:

$$V_{1a} \triangleq V_1 + \sum_{i=1}^p F_i \bar{M}_i F_i^T, \quad Q_a \triangleq V_{01} + QC^T + \sum_{i=1}^p D_i \bar{M}_i F_i^T,$$

$$D \triangleq \sum_{i=1}^p D_i \bar{M}_i D_i^T, \quad E \triangleq \sum_{i=1}^p E_i^T \bar{N}_i E_i, \quad A_Q \triangleq A - Q_a V_{1a}^{-1} C.$$

The following factorization lemma is needed for the statement of the main result. See [1] for details.

Lemma 8.1. If $\hat{Q}, \hat{P} \in \mathbb{N}^n$ and $\text{rank } \hat{Q}\hat{P} = n_e$, then there exist $n_e \times n$ G, Γ and $n_e \times n_e$ invertible M such that

$$\hat{Q}\hat{P} = G^T M \Gamma, \quad \Gamma G^T = I_{n_e}. \quad (8.1), (8.2)$$

Furthermore, G, M and Γ are unique except for a change of basis in \mathbb{R}^{n_e} .

Since $\hat{Q}\hat{P}$ is diagonalizable it has a group generalized inverse $(\hat{Q}\hat{P})^\# = G^T M^{-1} \Gamma$ and

$$\tau \triangleq \hat{Q}\hat{P}(\hat{Q}\hat{P})^\# = G^T \Gamma \quad (8.3)$$

is an oblique projection. Define the complementary projection $\tau_\perp \triangleq I_n - \tau$ and call (G, M, Γ) satisfying (8.1) and (8.2) a *projective factorization* of $\hat{Q}\hat{P}$.

Theorem 8.1. $(\mathcal{Q}, A_e, B_e, C_e, D_e) \in \mathcal{S}$ is an extremal of the Auxiliary Minimization problem with \mathcal{U} given by (5.1) if and only if there exist $Q, \hat{Q}, \hat{P} \in \mathbb{N}^n$ such that $\mathcal{Q}, A_e, B_e, C_e, D_e$ are given by

$$\mathcal{Q} = \begin{bmatrix} Q + \hat{Q} & \hat{Q}\Gamma^T \\ \Gamma\hat{Q} & \Gamma\hat{Q}\Gamma^T \end{bmatrix}, \quad (8.4)$$

$$A_e = \Gamma(A - Q_a V_{1a}^{-1} C + QE)G^T, \quad (8.5)$$

$$B_e = \Gamma Q_a V_{1a}^{-1}, \quad (8.6)$$

$$C_e = L\tau_{1\perp} G^T, \quad (8.7)$$

$$D_e = LQ\hat{C}^T(\hat{C}Q\hat{C}^T)^{-1}, \quad (8.8)$$

for some projective factorization (G, M, Γ) of $\hat{Q}\hat{P}$, and such that Q, \hat{Q}, \hat{P} satisfy

$$0 = AQ + QA^T + V_0 + D + QEQ - Q_a V_{1a}^{-1} Q_a^T + \tau_{1\perp} Q_a V_{1a}^{-1} Q_a^T \tau_{1\perp}^T, \quad (8.9)$$

$$0 = (A + QE)\hat{Q} + \hat{Q}(A + QE)^T + \hat{Q}E\hat{Q} + Q_a V_{1a}^{-1} Q_a^T - \tau_{1\perp} Q_a V_{1a}^{-1} Q_a^T \tau_{1\perp}^T, \quad (8.10)$$

$$0 = (A_Q + QE)^T \hat{P} + \hat{P}(A_Q + QE) + \tau_{1\perp}^T L^T R L \tau_{1\perp} - \tau_{1\perp}^T \tau_{1\perp}^T L^T R L \tau_{1\perp} \tau_{1\perp}, \quad (8.11)$$

$$\text{rank } \hat{Q} = \text{rank } \hat{P} = \text{rank } \hat{Q}\hat{P} = n_e. \quad (8.12)$$

where

$$\tau_1 \triangleq Q\hat{C}^T(\hat{C}Q\hat{C}^T)^{-1}\hat{C}, \quad \tau_{1\perp} \triangleq I_n - \tau_1. \quad (8.13)$$

Theorem 8.1 (proved in Appendix A) presents necessary conditions for the Auxiliary Minimization Problem which explicitly characterize extremals (\mathcal{Q} , A_e , B_e , C_e , D_e). These necessary conditions consist of a system of two modified Lyapunov equations and one modified Riccati equation coupled by two oblique projections τ and τ_1 and uncertainty terms. The projections τ and τ_1 correspond to reduced estimator order and singular observation noise, respectively.

Several special cases can immediately be discerned. For example, in the full-order estimator case $n_e = n$, set $\tau = I_n$ so that $\tau_{1\perp} = 0$. Now the last term in each of (8.9)–(8.11) can be deleted and G and Γ in (8.4)–(8.7) can be taken to be the identity. Furthermore, since \hat{Q} and \hat{P} now play no role in determining the optimal estimator, equations (8.10) and (8.11) are superfluous. If, furthermore, D_e , E_e , and F_e are zero, then (8.9) reduces to the standard observer Riccati equation of steady-state Kalman filter theory. Alternatively, the case in which the static estimator gain D_e is absent can be handled by ignoring (8.8) and setting $\tau_1 = 0$. If, furthermore, the uncertainty terms are deleted then the results of [1] are recovered.

9. Sufficient conditions for robust, reduced-order estimation

The main result guaranteeing robust estimation can now be stated.

Theorem 9.1. Suppose there exist \hat{Q} , $\hat{P} \in \mathbb{N}^n$ satisfying (8.9)–(8.12), let A_e , B_e , C_e , D_e be given by (8.5)–(8.8), and suppose that $(\tilde{V}^{1/2}, \tilde{A} + \Delta\tilde{A})$ is detectable for all $(\Delta A, \Delta C) \in \mathcal{U}$ with \mathcal{U} given by (5.1). Then A_e is asymptotically stable, $A + \Delta A$ is asymptotically stable for all $(\Delta A, \Delta C) \in \mathcal{U}$, and the estimation error satisfies the bound

$$J(A_e, B_e, C_e, D_e) \leq \text{tr } Q\tau_{1\perp}^T L^T R L \tau_{1\perp}. \quad (9.1)$$

Proof. Theorem 8.1 implies \mathcal{Q} given by (8.4) satisfies (7.2). With the detectability assumption the result follows from Proposition 7.1. \square

Remark 9.1. Note that if $\hat{C} = L$ then $C_e = 0$ and the estimation bound (9.1) is zero since $\hat{C}\tau_{1\perp} = 0$. This is, of course, to be expected since perfect estimation is achievable in this case.

Remark 9.2. The problem of designing reduced-order, robust estimators for unstable systems remains an area for future research.

Appendix A: Proof of Theorem 8.1

Partition $\tilde{n} \times \tilde{n}$ \mathcal{Q} , \mathcal{P} into $n \times n$, $n \times n_e$, and $n_e \times n_e$ subblocks as

$$\mathcal{Q} = \begin{bmatrix} Q_1 & Q_{12} \\ Q_{12}^T & Q_2 \end{bmatrix}, \quad \mathcal{P} = \begin{bmatrix} P_1 & P_{12} \\ P_{12}^T & P_2 \end{bmatrix},$$

and define the $n \times n$ nonnegative-definite matrices

$$\begin{aligned} Q &\triangleq Q_1 - Q_{12}Q_2^{-1}Q_{12}^T, & P &\triangleq P_1 - P_{12}P_2^{-1}P_{12}^T, \\ \hat{Q} &\triangleq Q_{12}Q_2^{-1}Q_{12}^T, & \hat{P} &\triangleq P_{12}P_2^{-1}P_{12}^T, \end{aligned}$$

and the $n_e \times n$, $n_e \times n_e$, $n_e \times n$ matrices

$$G \triangleq Q_2^{-1}Q_{12}^T, \quad M \triangleq Q_2P_2, \quad \Gamma \triangleq -P_2^{-1}P_{12}^T.$$

The existence of Q_2^{-1} and P_2^{-1} is shown below.

To optimize (7.1) over the open set \mathcal{S}' , where $\mathcal{S}' \triangleq \{(\mathcal{Q}, A_e, B_e, C_e, D_e) \in \mathcal{S} : (7.3) \text{ is satisfied}\}$, subject to the constraint (7.2), form the Lagrangian

$$\mathcal{L}(\mathcal{Q}, A_e, B_e, C_e, D_e) \triangleq \text{tr} \left[\lambda \mathcal{Q} \tilde{R} + \left(\tilde{A} \mathcal{Q} + \mathcal{Q} \tilde{A}^T + \sum_{i=1}^p \tilde{D}_i \tilde{M}_i \tilde{D}_i^T + \mathcal{Q} \tilde{E}_i^T \tilde{N}_i \tilde{E}_i \mathcal{Q} + \tilde{V} \right) \mathcal{P} \right],$$

where the multipliers $\lambda \geq 0$ and $\mathcal{P} \in \mathbb{R}^{\tilde{n} \times \tilde{n}}$ are not both zero. We thus obtain

$$\frac{\partial \mathcal{L}}{\partial \mathcal{Q}} = \tilde{A}^T \mathcal{P} + \mathcal{P} \tilde{A} + \sum_{i=1}^p \tilde{E}_i^T \tilde{N}_i \tilde{E}_i \mathcal{Q} \mathcal{P} + \mathcal{P} \mathcal{Q} \tilde{E}_i^T \tilde{N}_i \tilde{E}_i + \lambda \tilde{R}.$$

Setting $\partial \mathcal{L} / \partial \mathcal{Q} = 0$ yields ('vec' is defined in [3])

$$\tilde{\mathcal{A}}^T \text{vec } \mathcal{P} = -\lambda \text{vec } \tilde{R}.$$

Since $\tilde{\mathcal{A}}$ is assumed to be invertible, $\lambda = 0$ implies $\mathcal{P} = 0$. Hence, without loss of generality, set $\lambda = 1$. Since, furthermore, $\tilde{\mathcal{A}}$ is assumed to be asymptotically stable, \mathcal{P} is nonnegative definite. The stationarity conditions are given by

$$\frac{\partial \mathcal{L}}{\partial \mathcal{P}} = \tilde{A} \mathcal{Q} + \mathcal{Q} \tilde{A}^T + \sum_{i=1}^p \tilde{D}_i \tilde{M}_i \tilde{D}_i^T + \mathcal{Q} \tilde{E}_i^T \tilde{N}_i \tilde{E}_i \mathcal{Q} + \tilde{V} = 0, \quad (\text{A.1})$$

$$\frac{\partial \mathcal{L}}{\partial \mathcal{Q}} = \tilde{A}^T \mathcal{P} + \mathcal{P} \tilde{A} + \sum_{i=1}^p \tilde{E}_i^T \tilde{N}_i \tilde{E}_i \mathcal{Q} \mathcal{P} + \mathcal{P} \mathcal{Q} \tilde{E}_i^T \tilde{N}_i \tilde{E}_i + \tilde{R} = 0. \quad (\text{A.2})$$

$$\frac{\partial \mathcal{L}}{\partial A_e} = P_{12}^T Q_{12} + P_2 Q_2 = 0, \quad (\text{A.3})$$

$$\frac{\partial \mathcal{L}}{\partial B_e} = P_{12}^T V_{01} + (P_{12}^T Q_1 + P_2 Q_{12}^T) C^T + P_2 B_e V_{1a} = 0, \quad (\text{A.4})$$

$$\frac{\partial \mathcal{L}}{\partial C_e} = -RLQ_{12} + RD_e C Q_{12} + RC_e Q_2 = 0, \quad (\text{A.5})$$

$$\frac{\partial \mathcal{L}}{\partial D_e} = -RLQ_1 \hat{C}^T + RD_e \hat{C} Q_1 \hat{C}^T + RC_e Q_{12}^T \hat{C}^T = 0. \quad (\text{A.6})$$

Expanding (A.1) and (A.2) yields

$$0 = A Q_1 + Q_1 A^T + D + Q_1 E Q_1 + V_0, \quad (\text{A.7})$$

$$0 = A Q_{12} + Q_1 C^T B_e^T + Q_{12} A_e^T + Q_1 E Q_{12} + V_{01} B_e^T, \quad (\text{A.8})$$

$$0 = B_e C Q_{12} + A_e Q_2 + Q_{12}^T C^T B_e^T + Q_2 A_e^T + B_e V_{1a} B_e^T + Q_{12}^T E Q_{12}, \quad (\text{A.9})$$

$$0 = P_{12} A_e + A^T P_{12} + C^T B_e^T P_2 + E (P_{12}^T Q_1 + P_2 Q_{12}^T)^T - L^T R C_e + \hat{C}^T D_e^T R C_e, \quad (\text{A.10})$$

$$0 = P_2 A_e + A_e^T P_2 + C_e^T R C_e. \quad (\text{A.11})$$

Note that the (1,1) subblock of equation (A.2) characterizing P_1 has been omitted from the above equations since the estimator gains are independent of P_1 . Writing (A.9) as (see [1,9])

$$0 = (A_e + B_e C Q_{12} Q_2^+) Q_2 + Q_2 (A_e + B_e C Q_{12} Q_2^+)^T + Q_2 (Q_{12} Q_2^+)^T E Q_{12} Q_2^+ Q_2 + B_e V_{1a} B_e^T$$

where Q_2^+ is the Moore-Penrose or Drazin generalized inverse of Q_2 , it follows from [17], Lemmas 2.1 and 12.2, that Q_2 is positive definite. Similarly, (A.11) implies that P_2 is positive definite.

Next (8.4), (8.6)–(8.8) follow from the definition of \mathcal{Q} , (A.4)–(A.6) by using the identities

$$Q_1 = Q + \hat{Q}, \quad P_1 = P + \hat{P}, \quad Q_{12} = \hat{Q} \Gamma^T, \quad P_{12} = -\hat{P} G^T, \quad Q_2 = \Gamma \hat{Q} \Gamma^T, \quad P_2 = G \hat{P} G^T.$$

Computing either $\Gamma(A.8) - (A.9)$ or $G(A.10) + (A.11)$ yields (8.5). Inserting (8.5), (8.8) into (A.7), (A.11) and using (A.7) + $G^T \Gamma(A.8) - (A.8)G - (A.8G)^T$ and $G^T \Gamma(A.8) - (A.8)G - (A.8G)^T$ yields (8.9) and (8.10). Similarly, $\Gamma^T G(A.10) - (A.10)G - (A.10G)^T$ yields (8.11).

Finally, the proof can be reversed so that (8.5)–(8.12) yield (A.1)–(A.6) and (7.2). See [9] for details. \square

References

- [1] D.S. Bernstein and D.C. Hyland, The optimal projection equations for reduced-order state estimation, *IEEE Trans. Automat. Control* **30** (1985) 583–585.
- [2] S.P. Bhattacharya, The structure of robust observers, *IEEE Trans. Automat. Control* **31** (1976) 561–566.
- [3] J.W. Brewer, Kronecker products and matrix calculus in system theory, *IEEE Trans. Circuits and Systems* **29** (1978) 767–771.
- [4] S.S.L. Chang and T.K.C. Peng, Adaptive guaranteed cost control of systems with uncertain parameters, *IEEE Trans. Automat. Control* **17** (1972) 474–483.
- [5] R.C. Chung and P.R. Belanger, Minimum sensitivity filter for linear time-varying stochastic systems with uncertain parameters, *IEEE Trans. Automat. Control* **31** (1976) 98–100.
- [6] J.A. D'Appolito and C.E. Hutchinson, Low sensitivity filters for state estimation in the presence of large parameter uncertainties, *IEEE Trans. Automat. Control* **34** (1989) 310–312.
- [7] T.E. Djalens, Robust observers for systems with parameters, *Systems Control Lett.* **6** (1986) 367–368.
- [8] K. Furuta, S. Hara and S. Mori, A class of systems with the same observer, *IEEE Trans. Automat. Control* **31** (1976) 777–778.
- [9] W.M. Haddad, Robust Optimal Projection Control System Synthesis, Ph.D. Dissertation, Department of Mechanical Engineering, Florida Institute of Technology, Melbourne, FL (March 1987).
- [10] D. Hinrichsen and A.J. Pritchard, Stability radius for structured perturbations and the algebraic Riccati equation, *Systems Control Lett.* **8** (1987) 105–113.
- [11] B.N. Jain, Guaranteed error estimation in uncertain systems, *IEEE Trans. Automat. Control* **30** (1975) 730–732.
- [12] I.R. Petersen and C.V. Hollot, A Riccati equation approach to the stabilization of uncertain systems, *Automatica* **22** (1986) 397–411.
- [13] I.R. Petersen, A stabilization algorithm for a class of uncertain linear systems, *Systems Control Lett.* **8** (1987) 357–367.
- [14] R.T. Stefani, Reducing the sensitivity to parameter variations of a minimum-order reduced-order observer, *Internat. J. Control* **35** (1982) 983–995.
- [15] M. Toda and R.V. Patel, Bounds on estimation errors of discrete time filters under modelling uncertainty, *IEEE Trans. Automat. Control* **25** (1980) 1115–1121.
- [16] A. Vinkler and L.J. Wood, Multistep guaranteed cost control of linear systems with uncertain parameters, *J. Guidance Control* **2** (1979) 449–456.
- [17] W.M. Wonham, *Linear Multivariable Control: A Geometric Approach*, Springer, New York, 1979.

Lyapunov bound suggested by recent work of Petersen and Hollot is utilized in conjunction with the guaranteed cost approach of Chang and Peng to guarantee robust stability with robust performance bound. Necessary conditions which generalize the optimal projection equations for fixed-order dynamic compensation are used to characterize the controller which minimizes the performance bound. The design equations thus effectively serve as sufficient conditions for synthesizing dynamic output-feedback controllers which provide robust stability and performance.

1 INTRODUCTION

As is well known, LQR and LQG controllers lack guaranteed robustness with respect to arbitrary parameter variations [1], [2]. Thus it is not surprising that there is considerable interest in the analysis and synthesis of feedback controllers which are robust with respect to structured real-valued plant parameter uncertainty. The present paper was motivated in particular by the guaranteed cost control approach of Chang and Peng [3], [4] and the robust stability technique of Petersen and Hollot [5], [7]. In [3], Chang and Peng consider a modified Riccati equation whose solutions are guaranteed to provide both robust stability and performance over a specified range of parameter variations. On the other hand, Petersen and Hollot in [5], [7] consider a different modified Riccati equation which utilizes a quadratic Lyapunov bound to provide robust stability over a range of structured plant variations. In the present paper, we combine aspects of both of these approaches to obtain both robust stability and performance.

Our preference for the Petersen-Hollot bound over the bound originally proposed by Chang and Peng is based upon the fact that the former is differentiable with respect to the Riccati solution, while the latter is not. We exploit this smoothness by utilizing the optimal projection approach for fixed-order dynamic compensation [8] in place of full-state feedback considered in [3], [4], [6], [7]. A systematic, in-depth treatment of the Chang-Peng, Petersen-Hollot, and other bounds (such as the right shift/multiplicative white noise bound considered in [9]-[11]) will be the subject of a future paper [12].

As discussed in [8], the optimal projection approach to fixed-order dynamic compensation is based upon a system of two modified algebraic Riccati equations and two modified algebraic Lyapunov equations which directly generalize LQG theory to the case of reduced-order controllers. To ensure robust stability and performance for reduced-order controllers, the present paper utilizes the Petersen-Hollot quadratic Lyapunov technique to bound the performance of controllers of fixed dimension. The performance bound is then interpreted as the cost functional for an auxiliary minimization problem whose optimality conditions directly generalize the results of [8]. Specifically, we again obtain a coupled system of algebraic Riccati and Lyapunov equations with additional terms arising from the Petersen-Hollot bound. When uncertainty is absent, these equations specialize immediately to the result of [8] which, in turn, specializes to LQG when the compensator order is equal to the plant dimension.

Although the optimal projection equations are necessary conditions for optimality, it is important to stress that in the present paper they are obtained not for the original cost function, but rather for a bound on the cost. The necessary conditions for the auxiliary minimization problem thus effectively serve as sufficient conditions for the original problem. Hence, even if a numerical solution of the extended optimal projection equations fails to produce the globally optimal controller, robust stability and performance are still guaranteed for all local extremals. Our approach thus seeks to rectify one of the main drawbacks of necessity theory by guaranteeing both robust stability and performance. Nevertheless, a numerical algorithm for computing the global optimum is given in [15].

In summary, the main contribution of the present paper is the generalization of the optimal projection equations by means of the Petersen-Hollot quadratic Lyapunov bound to synthesize robustly stabilizing fixed-order dynamic compensators with guaranteed performance bound. It is interesting to note that even in the full-order case, our results, which specialize to a coupled system of three matrix equations, are distinct from the results of [5] which involve a pair of modified Riccati equations and an auxiliary inequality. Furthermore, the

The Optimal Projection Equations with Petersen-Hollot Bounds: Robust Stability and Performance Via Fixed-Order Dynamic Compensation for Systems with Structured Real-Valued Parameter Uncertainty

DENNIS S. BERNSTEIN AND WASSIM M. HADDAD

Dedicated to the memory of Vicki Ringer Jones
October 26, 1951-December 12, 1987

Abstract—A feedback control-design problem involving structured real-valued plant parameter uncertainties is considered. A quadratic

Manuscript received December 29, 1986; revised September 17, 1987. Paper recommended by Associate Editor, D. P. Lofte. The work of D. S. Bernstein was supported in part by the U.S. Air Force Office of Scientific Research under Contract F49620-86-C-0002.

D. S. Bernstein is with the Government Aerospace Systems Division, Harris Corporation, Melbourne, FL 32902.

W. M. Haddad is with the Department of Mechanical Engineering, Florida Institute of Technology, Melbourne, FL 32901.
IEEE Log Number 8718561

present paper provides a robust performance bound not obtained in [5]-[7]. An additional, conceptual benefit of our approach is a rigorous optimization interpretation for the Petersen-Holot Riccati equation approach. Finally, as shown in [20] for full-state feedback, the results given herein can be directly applied to the H_∞ design problem. For details, see [21].

Due to space constraints, the contents of the paper will not be reviewed here. We note only that the proof of Theorem 8.1, which has been omitted for this reason, can be found in [13], [14]. Finally, although numerical algorithms are outside the scope of this note, related results can be found in [15].

II. NOTATION AND DEFINITIONS

Note: All matrices have real entries.

| | |
|---|--|
| $\mathbb{R}, \mathbb{R}^{n \times n}, \mathbb{R}^n, \mathbb{E}$ | Real numbers, $r \times s$ real matrices, $\mathbb{R}^{n \times 1}$, expected value. |
| $I_r, (\cdot)^T$ | $r \times r$ identity matrix, transpose |
| $\mathbb{S}^r, \mathbb{H}^r, \mathbb{P}^r$ | $r \times r$ symmetric, nonnegative-definite, positive-definite matrices. |
| $Z_1 \leq Z_2, Z_1 < Z_2$ | $Z_2 - Z_1 \in \mathbb{H}^r, Z_2 - Z_1 \in \mathbb{P}^r, Z_1, Z_2 \in \mathbb{S}^r$. |
| $n, m, l, n_c; \bar{n}$ | Positive integers; $n + n_c$. |
| x, u, y, x_c, \bar{x} | n, m, l, n_c, \bar{n} -dimensional vectors. |
| $A, \Delta A; B, \Delta B; C, \Delta C$ | $n \times n$ matrices; $n \times m$ matrices; $l \times n$ matrices. |
| A_c, B_c, C_c | $n_c \times n_c; n_c \times l; m \times n_c$ matrices. |
| $\bar{A}, \Delta \bar{A}$ | $\begin{bmatrix} A & BC_c \\ B_c C & A_c \end{bmatrix}, \begin{bmatrix} \Delta A & \Delta BC_c \\ B_c \Delta C & 0 \end{bmatrix}$. |
| R_1, R_2 | $n \times n, m \times m$ state, control weighting matrices. $R_1 \geq 0, R_2 > 0$. |
| R_{12} | $n \times m$ cross weighting matrix; $R_1 - R_{12}R_2^{-1}R_{12}^T \geq 0$. |
| $w_1(\cdot), w_2(\cdot)$ | n, l -dimensional white noise. |
| V_1, V_2 | Intensity of $w_1(\cdot), w_2(\cdot)$; $V_1 \geq 0, V_2 > 0$. |
| V_{12} | $n \times l$ cross intensity of $w_1(\cdot), w_2(\cdot)$. |
| $\bar{w}(\cdot), \bar{V}$ | $\begin{bmatrix} w_1(\cdot) \\ B, w_2(\cdot) \end{bmatrix}, \begin{bmatrix} V_1 & V_{12}B_c^T \\ B, V_{12}^T & B, V_2 B_c^T \end{bmatrix}$ |
| \bar{R} | $\begin{bmatrix} R_1 & R_{12}C_c \\ C_c^T R_{12}^T & C_c^T R_2 C_c \end{bmatrix}$. |

III. ROBUST STABILITY AND ROBUST PERFORMANCE PROBLEMS

Let $\mathcal{U} \subset \mathbb{R}^{n \times n} \times \mathbb{R}^{n \times m} \times \mathbb{R}^{m \times n}$ denote the set of uncertain perturbations $(\Delta A, \Delta B, \Delta C)$ of the nominal plant matrices A, B , and C .

Robust Stability Problem: For fixed $n_c \leq n$, determine (A_c, B_c, C_c) such that the closed-loop system consisting of the n th-order controlled plant

$$\dot{x}(t) = (A + \Delta A)x(t) + (B + \Delta B)u(t), \quad t \in [0, \infty), \quad (3.1)$$

measurements

$$y(t) = (C + \Delta C)x(t), \quad (3.2)$$

and n_c th-order dynamic compensator

$$\dot{x}_c(t) = A_c x_c(t) + B_c y(t), \quad (3.3)$$

$$u(t) = C_c x_c(t) \quad (3.4)$$

is asymptotically stable for all $(\Delta A, \Delta B, \Delta C) \in \mathcal{U}$.

Robust Performance Problem: For fixed $n_c \leq n$, determine (A_c, B_c, C_c) such that, for the closed-loop system consisting of the n th-order disturbed plant

$$\dot{x}(t) = (A + \Delta A)x(t) + (B + \Delta B)u(t) + w_1(t), \quad t \in [0, \infty), \quad (3.5)$$

noisy measurements

$$y(t) = (C + \Delta C)x(t) + w_2(t), \quad (3.6)$$

and n_c th-order dynamic compensator (3.3), (3.4), the performance criterion

$$J(A_c, B_c, C_c) \triangleq \sup_{(\Delta A, \Delta B, \Delta C) \in \mathcal{U}} \limsup_{t \rightarrow \infty} \mathbb{E}\{x^T(t)R_1 x(t) + 2x^T(t)R_{12}u(t) + u^T(t)R_2 u(t)\} \quad (3.7)$$

is minimized.

Remark 3.1: Note that (3.7) is precisely the LQG criterion except for the supremum over \mathcal{U} for worst-case performance.

For each controller (A_c, B_c, C_c) and plant variation $(\Delta A, \Delta B, \Delta C) \in \mathcal{U}$, the undisturbed closed-loop system (3.1)-(3.4) is given by

$$\dot{\bar{x}}(t) = (\bar{A} + \Delta \bar{A})\bar{x}(t), \quad t \in [0, \infty), \quad (3.8)$$

while the disturbed closed-loop system (3.3)-(3.6) is

$$\dot{\bar{x}}(t) = (\bar{A} + \Delta \bar{A})\bar{x}(t) + \bar{w}(t), \quad t \in [0, \infty), \quad (3.9)$$

where $\bar{x}(t) \triangleq [x^T(t), x_c^T(t)]^T$ and $\bar{w}(\cdot)$ is white noise with intensity $\bar{V} \in \mathbb{H}^{\bar{n}}$.

IV. SUFFICIENT CONDITIONS FOR ROBUST STABILITY AND PERFORMANCE

In practice, steady-state performance is only of interest when the closed-loop system (3.8) is stable over \mathcal{U} . The following result expresses the performance in terms of the steady-state closed-loop second-moment matrix.

Lemma 4.1: Suppose (3.8) is stable for all $(\Delta A, \Delta B, \Delta C) \in \mathcal{U}$. Then

$$J(A_c, B_c, C_c) = \sup_{(\Delta A, \Delta B, \Delta C) \in \mathcal{U}} \text{tr } \bar{Q}_{\Delta A} \bar{R}, \quad (4.1)$$

where $\bar{Q}_{\Delta A} \triangleq \lim_{t \rightarrow \infty} \mathbb{E}\{\bar{x}(t)\bar{x}^T(t)\} \in \mathbb{H}^{\bar{n}}$ is the unique solution to

$$0 = (\bar{A} + \Delta \bar{A})\bar{Q}_{\Delta A} + \bar{Q}_{\Delta A}(\bar{A} + \Delta \bar{A})^T + \bar{V}. \quad (4.2)$$

We now seek upper bounds for $J(A_c, B_c, C_c)$.

Theorem 4.1: Let $\Omega: \mathbb{H}^{\bar{n}} \times \mathbb{H}^{n \times n} \times \mathbb{H}^{m \times m} \rightarrow \mathbb{S}^{\bar{n}}$ be such that

$$\Delta \bar{A} \bar{Q} + \bar{Q} \Delta \bar{A}^T \leq \Omega(\bar{Q}, B_c, C_c), \quad (\Delta A, \Delta B, \Delta C) \in \mathcal{U}, (\bar{Q}, B_c, C_c) \in \mathbb{H}^{\bar{n}} \times \mathbb{H}^{n \times n} \times \mathbb{H}^{m \times m}, \quad (4.3)$$

and, for given (A_c, B_c, C_c) , suppose there exists $\bar{Q} \in \mathbb{H}^{\bar{n}}$ satisfying

$$0 = \bar{A} \bar{Q} + \bar{Q} \bar{A}^T + \Omega(\bar{Q}, B_c, C_c) + \bar{V}, \quad (4.4)$$

and suppose the pair $(\bar{V}^{1/2}, \bar{A} + \Delta \bar{A})$ is stabilizable for all $(\Delta A, \Delta B, \Delta C) \in \mathcal{U}$. Then $\bar{A} + \Delta \bar{A}$ is asymptotically stable for all $(\Delta A, \Delta B, \Delta C) \in \mathcal{U}$,

$$\bar{Q}_{\Delta A} \leq \bar{Q}, \quad (\Delta A, \Delta B, \Delta C) \in \mathcal{U}, \quad (4.5)$$

where $\bar{Q}_{\Delta A}$ satisfies (4.2), and

$$J(A_c, B_c, C_c) \leq \text{tr } \bar{Q} \bar{R}. \quad (4.6)$$

Proof: For all $(\Delta A, \Delta B, \Delta C) \in \mathcal{U}$, (4.4) is equivalent to

$$0 = (\bar{A} + \Delta \bar{A})\bar{Q} + \bar{Q}(\bar{A} + \Delta \bar{A})^T + \Psi(\bar{Q}, B_c, C_c, \Delta \bar{A}) + \bar{V}, \quad (4.7)$$

where

$$\Psi(\bar{Q}, B_c, C_c, \Delta \bar{A}) \triangleq \Omega(\bar{Q}, B_c, C_c) - (\Delta \bar{A} \bar{Q} + \bar{Q} \Delta \bar{A}^T).$$

Note that by (4.3), $\Psi(\bar{Q}, B_c, C_c, \Delta \bar{A}) \geq 0$ for all $(\Delta A, \Delta B, \Delta C) \in \mathcal{U}$. Since $(\bar{V}^{1/2}, \bar{A} + \Delta \bar{A})$ is stabilizable for all $(\Delta A, \Delta B, \Delta C) \in \mathcal{U}$, it follows from [16, Theorem 3.6] that $((\bar{V} + \Psi(\bar{Q}, B_c, C_c, \Delta \bar{A}))^{1/2}, \bar{A} + \Delta \bar{A})$ is stabilizable for all $(\Delta A, \Delta B, \Delta C) \in \mathcal{U}$. Hence, [16, Lemma 12.2] implies $\bar{A} + \Delta \bar{A}$ is asymptotically stable for all $(\Delta A, \Delta B, \Delta C) \in \mathcal{U}$.

u. Next, subtracting (4.2) from (4.7) yields

$$0 = (\tilde{A} + \Delta\tilde{A})(Q - \tilde{Q}_{\Delta}) + (Q - \tilde{Q}_{\Delta})(\tilde{A} + \Delta\tilde{A})^T + \Psi(Q, B_c, C_c, \Delta\tilde{A})$$

or, equivalently (since $\tilde{A} + \Delta\tilde{A}$ is asymptotically stable),

$$Q - \tilde{Q}_{\Delta} = \int_0^{\infty} e^{(\tilde{A} + \Delta\tilde{A})t} \Psi(Q, B_c, C_c, \Delta\tilde{A}) e^{(\tilde{A} + \Delta\tilde{A})^T t} dt \geq 0,$$

which implies (4.5). Finally, (4.5) and (4.1) yield (4.6). \square

V. UNCERTAINTY STRUCTURE

To obtain explicit expressions for (A_c, B_c, C_c) , we require that $\Delta B = 0$, $(\Delta A, \Delta C) \in \mathcal{U}$. Hence, for simplicity, we write $(\Delta A, \Delta C) \in \mathcal{U}$. The dual case $\Delta B \neq 0$ and $\Delta C = 0$ is treated in Section X. Thus, \mathcal{U} is assumed to be of the form

$$\mathcal{U} = \left\{ (\Delta A, \Delta C) \in \mathbb{R}^{n \times n} \times \mathbb{R}^{p \times n} : \Delta A = \sum_{i=1}^p D_i M_i N_i E_i, \right. \\ \left. \Delta C = \sum_{i=1}^p F_i M_i N_i E_i, \quad M_i M_i^T \leq \tilde{M}_i, \quad N_i^T N_i \leq \tilde{N}_i, \quad i = 1, \dots, p \right\}, \quad (5.1)$$

where, for $i = 1, \dots, p$: $D_i \in \mathbb{R}^{n \times r_i}$, $E_i \in \mathbb{R}^{n \times n}$, and $F_i \in \mathbb{R}^{p \times r_i}$ are fixed matrices denoting the structure of the uncertainty; $\tilde{M}_i \in \mathbb{R}^{r_i \times r_i}$ and $\tilde{N}_i \in \mathbb{R}^{n \times n}$ are given uncertainty bounds; and $M_i \in \mathbb{R}^{r_i \times n}$ and $N_i \in \mathbb{R}^{n \times r_i}$ are uncertain matrices. The closed-loop system thus has structured uncertainty of the form

$$\Delta\tilde{A} = \sum_{i=1}^p \tilde{D}_i M_i N_i \tilde{E}_i,$$

where

$$\tilde{D}_i \triangleq \begin{bmatrix} D_i \\ B_c F_i \end{bmatrix}, \quad \tilde{E}_i \triangleq [E_i \quad 0].$$

The special case $\tilde{M}_i = \mu_i^2 I_{r_i}$, $\tilde{N}_i = \nu_i^2 I_n$ is worth noting.

Proposition 5.1: Let $\mu_i, \nu_i \geq 0$, $i = 1, \dots, p$. Then $M_i M_i^T \leq \mu_i^2 I_{r_i}$ and $N_i^T N_i \leq \nu_i^2 I_n$ if and only if $\sigma_{\max}(M_i) \leq \mu_i$ and $\sigma_{\max}(N_i) \leq \nu_i$.

Remark 5.1: The form of \mathcal{U} given by (5.1) is directly related to the structured stability radius introduced by Hinrichsen and Pritchard [17], [18]. Specifically, let $p = 1$, $\tilde{M}_1 = \mu_1 I_{r_1}$, $r_1 = s_1$, and $N_1 = \tilde{N}_1 = I_{r_1}$.

VI. THE PETERSEN-HOLLOT BOUND

Given \mathcal{U} , we now specify the bound Ω satisfying (4.3). Note that because of $\Delta B = 0$, Ω is independent of C_c . Hence, we write $\Omega(Q, B_c)$ for $\Omega(Q, B_c, C_c)$.

Proposition 6.1: The function

$$\Omega(Q, B_c) \triangleq \sum_{i=1}^p \tilde{D}_i M_i D_i^T + Q E_i^T N_i E_i Q \quad (6.1)$$

satisfies (4.3) with \mathcal{U} given by (5.1).

Proof: For $i = 1, \dots, p$,

$$0 \leq [\tilde{D}_i M_i - Q E_i^T N_i^T] [\tilde{D}_i M_i - Q E_i^T N_i^T]^T \\ = \tilde{D}_i M_i M_i^T \tilde{D}_i^T + Q E_i^T N_i^T N_i E_i Q - (\tilde{D}_i M_i N_i E_i Q + Q E_i^T N_i^T M_i^T \tilde{D}_i^T) \\ \leq \tilde{D}_i \tilde{M}_i \tilde{D}_i^T + Q E_i^T \tilde{N}_i E_i Q - (\tilde{D}_i M_i N_i E_i Q + Q E_i^T N_i^T M_i^T \tilde{D}_i^T).$$

Summing over i yields (4.3). \square

Remark 6.1: The bound (6.1) was originally proposed by Petersen in [5] for unit-rank perturbations with scalar uncertain parameters. A more general treatment appears in [7]. Note that we absorb the epsilon used in [7] into D_i and E_i .

VII. THE AUXILIARY MINIMIZATION PROBLEM

To optimize robust performance while guaranteeing robust stability, we consider the following problem.

Auxiliary Minimization Problem: Determine (Q, A_c, B_c, C_c) which minimizes

$$j(Q, A_c, B_c, C_c) \triangleq \text{tr } Q \tilde{R} \quad (7.1)$$

subject to

$$Q \in \mathcal{H}^+, \quad (7.2)$$

$$0 = \tilde{A}Q + Q\tilde{A}^T + \sum_{i=1}^p [\tilde{D}_i M_i \tilde{D}_i^T + Q E_i^T N_i E_i Q] + \tilde{P}, \quad (7.3)$$

$$(\tilde{P}^{\frac{1}{2}}, \tilde{A} + \Delta\tilde{A}) \text{ is stabilizable, } (\Delta A, \Delta C) \in \mathcal{U}. \quad (7.4)$$

Proposition 7.1: If (Q, A_c, B_c, C_c) satisfies (7.2)–(7.4), then $\tilde{A} + \Delta\tilde{A}$ is asymptotically stable for all $(\Delta A, \Delta C) \in \mathcal{U}$ and

$$J(A_c, B_c, C_c) \leq j(Q, A_c, B_c, C_c). \quad (7.5)$$

Proof: With Ω given by (6.1), the hypotheses of Theorem 4.1 are satisfied so that robust stability is guaranteed with performance bound (4.6). \square

VIII. NECESSARY CONDITIONS FOR THE AUXILIARY MINIMIZATION PROBLEM

Rigorous derivation of the necessary conditions for the Auxiliary Minimization Problem requires additional technical assumptions. Specifically, in addition to (7.2), we restrict (Q, A_c, B_c, C_c) to the open set

$$\mathcal{S} \triangleq \{(Q, A_c, B_c, C_c) : Q \in \mathcal{H}^+, \tilde{Q} \text{ is asymptotically stable,}$$

and (A_c, B_c, C_c) is controllable and observable\},

where (see [19] for the definition of the Kronecker sum)

$$\tilde{Q} \triangleq \left(\tilde{A} + \sum_{i=1}^p Q E_i^T N_i E_i \right) \oplus \left(\tilde{A} + \sum_{i=1}^p Q E_i^T N_i E_i \right).$$

Furthermore, the constraint (7.4) will not be accounted for explicitly since it can be shown that the compactness of \mathcal{U} implies that the set of (A_c, B_c, C_c) satisfying (7.4) is open.

Remark 8.1: The constraint $(Q, A_c, B_c, C_c) \in \mathcal{S}$ is not required for either robust stability or robust performance since Proposition 7.1 shows that only (7.2)–(7.4) are needed. Rather, the set \mathcal{S} constitutes sufficient conditions under which the Lagrange multiplier technique is applicable to the Auxiliary Minimization Problem. Specifically, the condition $Q > 0$ replaces (7.2) by an open set constraint, the stability of \tilde{Q} serves as a normality condition, and (A_c, B_c, C_c) minimal is a nondegeneracy condition.

For arbitrary $Q, P \in \mathbb{R}^{n \times n}$ define the following notation:

$$D \triangleq \sum_{i=1}^p \tilde{D}_i M_i D_i^T, \quad E \triangleq \sum_{i=1}^p E_i^T N_i E_i,$$

$$P_s \triangleq B^T P + R_{12}^T, \quad Q_s \triangleq Q C^T + V_{12} + \sum_{i=1}^p \tilde{D}_i M_i F_i^T,$$

$$A_p \triangleq A - B R_{12}^{-1} P_s, \quad A_0 \triangleq A - Q_s V_{12}^{-1} C, \quad V_{12} \triangleq V_{12} + \sum_{i=1}^p F_i M_i F_i^T.$$

The following factorization lemma is needed. For details, see [8].

Lemma 8.1: If $\tilde{Q}, \tilde{P} \in \mathbb{H}^n$ and $\text{rank } \tilde{Q}\tilde{P} = n_c$, then there exist $n_c \times n$ G, Γ , and $n_c \times n_c$ invertible M such that

$$\tilde{Q}\tilde{P} = G^T M \Gamma, \quad (8.1)$$

$$\Gamma G^T = I_{n_c}. \quad (8.2)$$

Furthermore, G , M , and Γ are unique except for a change of basis in \mathbb{R}^n . As shown in [8], the matrix τ defined by

$$\tau \triangleq \hat{Q}\hat{P}(\hat{Q}\hat{P})^\dagger = G^\dagger \Gamma \quad (8.3)$$

is an oblique projection where $(\cdot)^\dagger$ denotes group generalized inverse [8]. For convenience, define the complementary projection $\tau_\perp \triangleq I_n - \tau$.

Theorem 8.1: If $(Q, A_c, B_c, C_c) \in \mathcal{S}$ solves the Auxiliary Minimization Problem with \mathcal{U} given by (5.1), then there exist $Q, P, \hat{Q}, \hat{P} \in \mathbb{R}^n$ such that

$$Q = \begin{bmatrix} Q + \hat{Q} & \hat{Q}\Gamma^\dagger \\ \Gamma\hat{Q} & \Gamma\hat{Q}\Gamma^\dagger \end{bmatrix}, \quad (8.4)$$

$$A_c = \Gamma(A - BR_2^{-1}P_c - Q_cV_{2c}^{-1}C + QE)G^\dagger, \quad (8.5)$$

$$B_c = \Gamma Q_c V_{2c}^{-1}, \quad (8.6)$$

$$C_c = -R_2^{-1}P_c G^\dagger, \quad (8.7)$$

and such that Q, P, \hat{Q}, \hat{P} satisfy

$$0 = AQ + QA^\dagger + V_1 + D + QE\hat{Q} - Q_cV_{2c}^{-1}Q_c^\dagger + \tau_\perp Q_cV_{2c}^{-1}Q_c^\dagger\tau_\perp, \quad (8.8)$$

$$0 = [A + (Q + \hat{Q})E]^\dagger P + P[A + (Q + \hat{Q})E] + R_1 - P_c^\dagger R_2^{-1}P_c + \tau_\perp^\dagger P_c^\dagger R_2^{-1}P_c\tau_\perp, \quad (8.9)$$

$$0 = (A_p + QE)\hat{Q} + \hat{Q}(A_p + QE)^\dagger + \hat{Q}E\hat{Q} + Q_cV_{2c}^{-1}Q_c^\dagger - \tau_\perp Q_cV_{2c}^{-1}Q_c^\dagger\tau_\perp, \quad (8.10)$$

$$0 = (A_0 + QE)^\dagger \hat{P} + \hat{P}(A_0 + QE) + P_c^\dagger R_2^{-1}P_c - \tau_\perp^\dagger P_c^\dagger R_2^{-1}P_c\tau_\perp, \quad (8.11)$$

$$\text{rank } \hat{Q} = \text{rank } \hat{P} = \text{rank } \hat{Q}\hat{P} = n_c. \quad (8.12)$$

Furthermore, the auxiliary cost is given by

$$J(Q, A_c, B_c, C_c) = \text{tr}[(Q + \hat{Q})R_1 + \hat{Q}PBR_2^{-1}P_c - R_{12}R_2^{-1}P_c\hat{Q}]. \quad (8.13)$$

Conversely, if there exist $Q, P, \hat{Q}, \hat{P} \in \mathbb{R}^n$ satisfying (8.8)–(8.12), then (Q, A_c, B_c, C_c) given by (8.4)–(8.7) satisfy (7.2) and (7.3) with cost (8.13).

Proof: See [13], [14]. \square

Remark 8.2: Theorem 8.1 presents necessary conditions for the Auxiliary Minimization Problem which explicitly characterize extremal quadruples (Q, A_c, B_c, C_c) . These necessary conditions consist of a system of two modified Riccati equations and two modified Lyapunov equations coupled by both the optimal projection τ and uncertainty terms. Several special cases can immediately be discerned. For example, in the full-order case $n_c = n$, set $\tau = I_n$ so that $\tau_\perp = 0$. Now the last term in each of (8.8)–(8.11) can be deleted and G and Γ in (8.5)–(8.7) can be taken to be the identity. Furthermore, \hat{P} plays no role so that (8.11) is superfluous. Note that in this case, (8.8) is independent of P and \hat{Q} . Setting further D, E , and F to zero, it can be seen that (8.10) and (8.11) drop out, while (8.8) and (8.9) reduce to the standard separated Riccati equations of LQG theory. If, alternatively, the reduced-order constraint is retained, but the uncertainty terms are deleted, then the results of [8] are recovered.

Remark 8.3: When solving (8.8)–(8.12) numerically, the uncertainty terms can be adjusted to examine tradeoffs between performance and robustness. Specifically, the bounds \bar{M} and \bar{N} , and structure matrices D, E , and F appearing in Q_c, D, E , and V_{2c} can be varied systematically to determine the region of solvability of (8.8)–(8.12).

IX. SUFFICIENT CONDITIONS FOR ROBUST STABILITY AND PERFORMANCE

Theorem 9.1: Suppose there exist $Q, P, \hat{Q}, \hat{P} \in \mathbb{R}^n$ satisfying (8.8)–(8.12), and assume that $(P^{1/2}, \bar{A} + \Delta\bar{A})$ is stabilizable for all $(\Delta A, \Delta C) \in \mathcal{U}$ with A_c, B_c, C_c given by (8.5)–(8.7) and \mathcal{U} given by (5.1). Then \bar{A}

+ $\Delta\bar{A}$ is asymptotically stable for all $(\Delta A, \Delta C) \in \mathcal{U}$ and the closed-loop performance is bounded by (8.13).

Proof: Theorem 8.1 implies that Q given by (8.4) satisfies (7.2) and (7.3). With the stabilizability assumption, the result follows from Proposition 7.1. \square

X. THE DUAL CASE

In place of (5.1), assume now that $\Delta C = 0, (\Delta A, \Delta B, \Delta C) \in \mathcal{U}$, and define

$$\mathcal{U} = \left\{ (\Delta A, \Delta B) \in \mathbb{R}^{n \times n} \times \mathbb{R}^{n \times m} : \Delta A = \sum_{i=1}^p D_i M_i N_i E_i, \right. \\ \left. \Delta B = \sum_{i=1}^p D_i M_i N_i G_i, \quad M_i M_i^\dagger \leq \bar{M}_i, \quad N_i^\dagger N_i \leq \bar{N}_i, \quad i = 1, \dots, p \right\}, \quad (10.1)$$

where, for $i = 1, \dots, p$: $D_i \in \mathbb{R}^{n \times n}$, $E_i \in \mathbb{R}^{n \times n}$, and $G_i \in \mathbb{R}^{n \times m}$ are fixed matrices denoting the structure of the uncertainty; and $\bar{M}_i, \bar{N}_i, M_i$, and N_i are as before. For arbitrary $Q, P \in \mathbb{R}^{n \times n}$ define the following notation:

$$\hat{P}_c \triangleq B^\dagger P + R_2^{-1} + \sum_{i=1}^p G_i^\dagger \bar{N}_i E_i, \quad \hat{Q}_c \triangleq Q C^\dagger + V_{12},$$

$$\hat{A}_p \triangleq A - BR_2^{-1}\hat{P}_c, \quad \hat{A}_0 \triangleq A - Q_cV_{2c}^{-1}C, \quad R_{2c} \triangleq R_2 + \sum_{i=1}^p G_i^\dagger \bar{N}_i G_i.$$

The main result guaranteeing robust stability and performance for the dual problem can now be stated. For details, see [13], [14].

Theorem 10.1: Suppose there exist $P, Q, \hat{P}, \hat{Q} \in \mathbb{R}^n$ satisfying (8.12) and

$$0 = A^\dagger P + PA + R_1 + E + PDP - \hat{P}_c^\dagger R_{2c}^{-1}\hat{P}_c + \tau_\perp^\dagger \hat{P}_c^\dagger R_{2c}^{-1}\hat{P}_c\tau_\perp, \quad (10.2)$$

$$0 = [A + D(P + \hat{P})]Q + Q[A + D(P + \hat{P})]^\dagger + V_1 - Q_cV_{2c}^{-1}\hat{Q}_c + \tau_\perp Q_cV_{2c}^{-1}\hat{Q}_c^\dagger\tau_\perp, \quad (10.3)$$

$$0 = (\hat{A}_0 + DP)^\dagger \hat{P} + \hat{P}(\hat{A}_0 + DP) + \hat{P}D\hat{P} + \hat{P}_c^\dagger R_{2c}^{-1}\hat{P}_c - \tau_\perp^\dagger \hat{P}_c^\dagger R_{2c}^{-1}\hat{P}_c\tau_\perp, \quad (10.4)$$

$$0 = (\hat{A}_p + DP)\hat{Q} + \hat{Q}(\hat{A}_p + DP)^\dagger + \hat{Q}_cV_{2c}^{-1}\hat{Q}_c^\dagger - \tau_\perp \hat{Q}_cV_{2c}^{-1}\hat{Q}_c^\dagger\tau_\perp, \quad (10.5)$$

and assume that $(\bar{R}^{1/2}, \bar{A} + \Delta\bar{A})$ is detectable for all $(\Delta A, \Delta B) \in \mathcal{U}$ with A_c, B_c, C_c given by

$$A_c = \Gamma(A - Q_cV_{2c}^{-1}C - BR_2^{-1}\hat{P}_c + DP)G^\dagger, \quad (10.6)$$

$$B_c = \Gamma Q_cV_{2c}^{-1}, \quad (10.7)$$

$$C_c = -R_{2c}^{-1}\hat{P}_c G^\dagger, \quad (10.8)$$

and \mathcal{U} given by (10.1). Then, with (10.6)–(10.8), $\bar{A} + \Delta\bar{A}$ is asymptotically stable for all $(\Delta A, \Delta B) \in \mathcal{U}$ and the performance of the closed-loop system satisfies

$$J(A_c, B_c, C_c) \leq \text{tr}[(P + \hat{P})V_1 + \hat{Q}_cV_{2c}^{-1}CQ\hat{P} - \hat{P}_cQ_cV_{2c}^{-1}V_{12}]. \quad (10.9)$$

Remark 10.1: Even in the case $\Delta B = 0, \Delta C = 0$, the performance bounds (8.13) and (10.9) are generally different.

Remark 10.2: The case in which ΔB and ΔC are simultaneously nonzero also appears to be tractable and leads to additional terms in the design equations. The bound considered in [11] also permits this case.

REFERENCES

- [1] J. C. Doyle, "Guaranteed margins for LQG regulators," *IEEE Trans. Automat. Contr.*, vol. AC-23, pp. 756–757, 1978.

- [2] E. Soroka and U. Shaked, "On the robustness of LQ regulators," *IEEE Trans. Automat. Contr.*, vol. AC-29, pp. 664-665, 1984.
- [3] S. S. L. Chang and T. K. C. Peng, "Adaptive guaranteed cost control of systems with uncertain parameters," *IEEE Trans. Automat. Contr.*, vol. AC-17, pp. 474-483, 1972.
- [4] A. Vinkler and L. J. Wood, "Multistep guaranteed cost control of linear systems with uncertain parameters," *J. Guidance Contr.*, vol. 2, pp. 449-456, 1979.
- [5] I. R. Petersen, "A Riccati equation approach to the design of stabilizing controllers and observers for a class of uncertain linear systems," in *Proc. Amer. Contr. Conf.*, Boston, MA, June 1985, pp. 772-777.
- [6] I. R. Petersen and C. V. Hollot, "A Riccati equation approach to the stabilization of uncertain systems," *Automatica*, vol. 22, pp. 397-411, 1986.
- [7] I. R. Petersen, "A stabilization algorithm for a class of uncertain linear systems," *Syst. Contr. Lett.*, vol. 8, pp. 351-357, 1987.
- [8] D. C. Hyland and D. S. Bernstein, "The optimal projection equations for fixed-order dynamic compensation," *IEEE Trans. Automat. Contr.*, vol. AC-29, pp. 1034-1037, 1984.
- [9] D. S. Bernstein and S. W. Greeley, "Robust controller synthesis using the maximum entropy design equations," *IEEE Trans. Automat. Contr.*, vol. AC-31, pp. 362-364, 1986.
- [10] —, "Robust output-feedback stabilization: Deterministic and stochastic perspectives," in *Proc. Amer. Contr. Conf.*, Seattle, WA, June 1986, pp. 1818-1826.
- [11] D. S. Bernstein, "Robust static and dynamic output feedback stabilization: Deterministic and stochastic perspectives," *IEEE Trans. Automat. Contr.*, vol. AC-32, pp. 1076-1084, 1987.
- [12] D. S. Bernstein and W. M. Haddad, "Robust stability and performance of linear dynamic systems via a priori quadratic Lyapunov bounds," in preparation.
- [13] W. M. Haddad, "Robust optimal projection control-system synthesis," Ph.D. dissertation, Dep. Mech. Eng., Florida Inst. Technol., Melbourne, FL, Mar. 1987.
- [14] D. S. Bernstein and W. M. Haddad, "The optimal projection equations with Petersen-Hollot bounds: Robust controller synthesis with guaranteed structured stability radius," in *Proc. IEEE Conf. Decision Contr.*, Los Angeles, CA, Dec. 1987, pp. 1308-1318.
- [15] S. Richter, "A homotopy algorithm for solving the optimal projection equations for fixed-order dynamic compensation: Existence, convergence and global optimality," in *Proc. Amer. Contr. Conf.*, Minneapolis, MN, June 1987, pp. 1527-1531.
- [16] W. M. Wonham, *Linear Multivariable Control: A Geometric Approach*. New York: Springer-Verlag, 1979.
- [17] D. Hinrichsen and A. J. Pritchard, "Stability radii of linear systems," *Syst. Contr. Lett.*, vol. 7, pp. 1-10, 1986.
- [18] D. Hinrichsen and A. J. Pritchard, "Stability radius for structured perturbations and the algebraic Riccati equation," *Syst. Contr. Lett.*, vol. 8, pp. 105-113, 1987.
- [19] J. W. Brewer, "Kronecker products and matrix calculus in system theory," *IEEE Trans. Circuits Syst.*, vol. CAS-25, pp. 772-781, 1978.
- [20] I. R. Petersen, "Disturbance attenuation and H^∞ optimization: A design method based on the algebraic Riccati equation," *IEEE Trans. Automat. Contr.*, vol. AC-32, pp. 427-429, 1987.
- [21] D. S. Bernstein and W. M. Haddad, "LQG control with an H_∞ performance bound," *Amer. Contr. Conf.*, June 1988, Atlanta, GA.

A Frequency Response-Based Model Order Selection Criterion

DAVID J. CLOUD AND BASIL KOUVARITAKIS

Abstract—The use of weighting sequence models to describe the dynamics of physical systems provides an effective means of translating the uncertainty associated with the model parameter estimates derived from noisy input/output data into corresponding frequency response uncertainty information. However, an appropriate truncation level must be established to accomplish this task. This paper addresses the truncation problem from a frequency response perspective and proposes a new criterion based on frequency response considerations to select the proper truncation.

Manuscript received January 6, 1987; revised May 18, 1987. This work was supported by the U.S. Air Force and the U. K. Science and Engineering Research Council (SERC). The authors are with the Department of Engineering Science, University of Oxford, Oxford OX1 3PJ, England.
IEEE Log Number 8718562.

I. INTRODUCTION

Recent work on robustness analysis methods for multivariable systems has led to the development of structured singular value techniques (e.g., [1], [2]) which provide a means of assessing the impact of system uncertainty on closed-loop stability and performance. A key missing element in these analysis methods is the ability to describe the frequency response uncertainty associated with any given system. When system identification techniques are used to derive the system description, however, it becomes possible to quantify system uncertainty statistically, and recent efforts have demonstrated that uncertainty information on the estimated parameters of the system model can be transformed into corresponding information on the frequency response uncertainty of the system [3], [4]. For difference equation models, the transformation from the parameter space to the frequency domain is nonlinear, a result which necessitates the use of linear approximations and produces a statistical description of uncertainty that fails to account for the interfrequency dependence of the frequency response estimates. On the other hand, Cloud and Kouvaritakis [4] have shown that these problems can be avoided by the use of weighting sequence models to describe system dynamics.

The uncertainty description developed in [4] assumes that the system can be accurately described by a finite weighting sequence model, an assumption that is valid for all stable systems. This assumption, in turn, implies that the "correct" model order (i.e., truncation level) is known. As a result, the identification process must not only be able to generate appropriate parameter estimates, it must also be able to identify the "correct" level of truncation. In effect, this second requirement is a restatement of the standard model order selection problem, a problem which has been widely investigated in the literature (e.g., [5]). But for the frequency response applications of interest here, appropriate solutions must focus on generating accurate frequency response information. When this perspective is taken, it becomes clear that the standard order selection criteria are not well suited to the task because they focus on generating accurate input/output descriptions rather than accurate frequency response descriptions.

In this paper, a new criterion is derived to identify the "correct" truncation level based on frequency response considerations. The development begins by highlighting a geometric interpretation of the standard "input/output" order selection problem. These geometric results are then transformed into the frequency domain to produce the new "frequency response-based" criterion for truncation selection, and simulation results are presented to demonstrate its use. Armed with the "correct" truncation level generated by this criterion, it is now possible to implement the techniques described in [4] to produce a valid description of frequency response uncertainty for any given system.

II. MODEL ORDER SELECTION: A GEOMETRIC PERSPECTIVE

Consider the discrete-time system with weighting sequence elements $\{\theta_1, \theta_2, \theta_3, \dots\}$ whose true response at sample k to the set of inputs $\{u(k-1), u(k-2), \dots, u(0)\}$ is given by

$$y^0(k) = \sum_{i=1}^{\infty} \theta_i u(k-i) = d'_k \theta^0 \quad (1)$$

where $\theta^0 = [\theta_1 \dots \theta_\infty]'$ and $d'_k = [u(k-1) \dots u(0)]$. The measured output at sample k is then given by

$$y^m(k) = y^0(k) + \epsilon(k) \quad (2)$$

where $\epsilon(k)$ is assumed to be an element of a white noise sequence with variance σ_ϵ^2 . For a set of N measurements, we may stack the scalars $y^m(k)$ and $\epsilon(k)$ as elements of the vectors y^m and ϵ , respectively, and may then rewrite (2) as

$$y^m = y^0 + \epsilon = D^0 \theta^0 + \epsilon \quad (3)$$

where the rows of D^0 are given by the vectors d'_k for $k = 1, \dots, N$.

Submitted June 1986
Reviews returned April 1987
Resubmitted May 1987
Reviews returned January 1988
Resubmitted February 1988

Robust Stability and Performance
via
Fixed-Order Dynamic Compensation

by

Dennis S. Bernstein
Harris Corporation
Melbourne, FL 32902

Abstract

Two robust control-design problems are considered. The Robust Stabilization Problem involves deterministically modeled bounded but unknown time-varying parameter variations, while the Robust Performance Problem includes, in addition, a quadratic performance criterion averaged over stochastic disturbances and maximized over the admissible parameter variations. For both problems the design goal is a fixed-order (i.e., reduced- or full-order) dynamic (strictly proper) feedback compensator. A sufficient condition for solving the Robust Stabilization Problem is given by means of a quadratic Lyapunov function parameterized by the compensator gains. For the Robust Performance Problem the Lyapunov function provides an upper bound for the closed-loop performance. This leads to consideration of the Auxiliary Minimization Problem: Minimize the performance bound over the class of fixed-order controllers subject to the Lyapunov-function constraint. Necessary conditions for optimality in the auxiliary problem thus serve as sufficient conditions for robust stability and performance in the original problem. Two particular bounds are considered for constructing the quadratic Lyapunov function. The first corresponds to a right shift/multiplicative white noise model, while the second was suggested by recent work of Petersen and Hollot. The main result is an extended version of the optimal projection equations for fixed-order dynamic compensation whose solutions are guaranteed to provide both robust stability and robust performance.

This research was supported in part by the Air Force Office of Scientific Research under contract F49620-86-C-0002.

1. Introduction

Although considerable effort has been devoted to frequency-domain robust control-design methods ([1-10]), there remain open questions concerning stability with respect to real-valued, structured plant parameter variations ([11-13]). Specifically, it is shown in [11-13] that classical gain and phase margin specifications can be satisfied while sensitivity to structured plant parameter variations can be arbitrarily large. From a time-domain point of view, the parametric robustness problem has been widely studied using Lyapunov's second method as the principal technique ([14-28]).

In the present paper we develop an approach to control design which provides sufficient conditions for robust stability and performance over a prescribed range of time-varying structured plant parameter variations by means of a feedback law in the form of a fixed-order (i.e., reduced- or full-order) dynamic (strictly proper) compensator. The approach is based upon the merging of two techniques, namely, the guaranteed cost control approach to robust performance ([14,17]) and the optimal projection approach to quadratically optimal fixed-order dynamic compensation ([29,30]). One of our goals is to obtain robust output-feedback compensators rather than full-state-feedback controllers. Also, since we wish to account for real-time computational burden in implementing the controller, we impose a constraint on the dimension (i.e., order) of the dynamic compensator. This approach thus generalizes standard LQG theory which yields full-order output-feedback controllers for systems without parameter uncertainty. We note that our approach is constructive in the sense that upon satisfaction of the sufficient conditions, the feedback gains required for implementing the robust feedback controller are explicitly synthesized. Existential issues are also addressed herein, although to a lesser extent. For further background, see [29,30]. For extensions to nonstrictly proper controllers see [31] and for extensions to H_∞ control see [32].

To explain the rationale behind the development we briefly describe the main elements of the approach. The following discussion is intended to be descriptive; precise conditions appear in the main body of the paper.

1. Robust Stability Problem. For a nominal linear time-invariant (A,B,C) system we consider deterministically modeled bounded but otherwise unknown Lebesgue measurable time-varying parameter variations of the form

$$A + \sum_{i=1}^p \hat{\sigma}_i(t) A_i, \quad B + \sum_{i=1}^p \hat{\sigma}_i(t) B_i, \quad C + \sum_{i=1}^p \hat{\sigma}_i(t) C_i. \quad (1.1)$$

The nominal matrices A, B, C and the perturbation matrices A_i, B_i, C_i denoting the structure of

the parametric uncertainty are assumed known, while the time-varying uncertain parameters $\hat{\sigma}_i(t)$ are assumed only to satisfy the bounds

$$|\hat{\sigma}_i(t)| \leq \delta_i, \quad i = 1, \dots, p, \quad t \in [0, \infty). \quad (1.2)$$

The form of (1.1) permits an arbitrary number of uncertain parameters with arbitrary linear structure. Although we do not require matching conditions as in [21], the linear structure of (1.1) is more restrictive than the functional form $A(q(t))$ used in [21]. It is this structure which we exploit to obtain sufficiency conditions. Note also that the representation (1.1) is independent of state space basis since replacing A by SAS^{-1} corresponds to replacing A_i by SA_iS^{-1} . As will be seen, our robustness bounds and optimality conditions are also basis independent. Also, scaling techniques ([6,7]) will not play a role here. Finally, we note that because of the time-varying nature of the uncertain perturbations (1.1) it is virtually impossible to determine the *actual* stability region of a given design by means of empirical methods.

2. Quadratic Lyapunov Function. As a sufficient condition for characterizing solutions of the Robust Stability Problem we consider a closed-loop quadratic Lyapunov function $V(\tilde{x}) = \tilde{x}^T P \tilde{x}$, where the matrix P satisfies

$$0 = \tilde{A}^T P + P \tilde{A} + \Omega(P, B_c, C_c) \quad (1.3)$$

and the function Ω is a bound satisfying

$$\sum_{i=1}^p \sigma_i (\tilde{A}_i^T P + P \tilde{A}_i) < \Omega(P, B_c, C_c) \quad (1.4)$$

over the parameter range

$$|\sigma_i| \leq \delta_i, \quad i = 1, \dots, p. \quad (1.5)$$

Note that the constant σ_i in (1.4) and (1.5) plays the role of $\hat{\sigma}_i(t)$, i.e., t is "frozen" in (1.4) and (1.5). In (1.3) and (1.4) \tilde{A} and \tilde{A}_i denote the closed-loop dynamics and closed-loop parameter-uncertainty matrices given by

$$\tilde{A} = \begin{bmatrix} A & BC_c \\ B_c C & A_c \end{bmatrix}, \quad \tilde{A}_i = \begin{bmatrix} A_i & B_i C_c \\ B_c C_i & 0 \end{bmatrix}. \quad (1.6)$$

Since \tilde{A}_i is independent of A_c , Ω depends only on B_c and C_c . As discussed later in this section, (1.4) is automatically satisfied by construction of the function Ω . Furthermore, the existence of a solution P to (1.3) need not be verified directly but is rather a result of numerically solving the optimality conditions discussed below.

3. Robust Performance Problem. In addition to the *deterministic* parameter uncertainty model (1.1), (1.2), the Robust Performance Problem includes *stochastic* plant disturbances and measurement noise with performance measured by means of the quadratic functional

$$\tilde{J}(t) = x^T(t)R_1x(t) + 2x^T(t)R_{12}u(t) + u^T(t)R_2u(t). \quad (1.7)$$

To obtain a steady-state design problem we 1) average $\tilde{J}(t)$ over the disturbance and measurement noise statistics; 2) pass to the steady-state limit; and 3) maximize over the class of parameter uncertainties. Hence the performance of a given controller (A_c, B_c, C_c) is given by

$$J(A_c, B_c, C_c) = \sup_{\theta(\cdot)} \limsup_{t \rightarrow \infty} \mathbb{E}[\tilde{J}(t)]. \quad (1.8)$$

The use of "limsup" is a technicality which account for cases in which the steady-state limit may not exist. Note that although (1.8) is an averaging criterion over the disturbances as in LQG theory, it is also a worst-case measure over the uncertain parameters. Thus (1.8) is a *hybrid* criterion in the sense that is *stochastic* in the disturbance space (i.e., external uncertainties) and *deterministic* in the parameter space (i.e., internal uncertainties). By "internal uncertainties" we have in mind quantities such as mass, damping or stiffness, and by "external uncertainties" we are referring to phenomena such as turbulent flow for which only power spectrum statistics may be available. No claim is made, however, with regard to the universal validity of such a mathematical uncertainty model. In particular applications, uncertainty models which are either wholly deterministic or wholly stochastic may be more appropriate. In general, our setting appears to be consistent with the available literature (see [1-28]).

4. Performance Bound. To obtain a tractable design problem the matrix P is used to bound the performance of each controller solving the Robust Stability Problem. Specifically, by assuming in addition to (1.4) that

$$\sum_{i=1}^p \sigma_i(\tilde{A}_i^T P + P \tilde{A}_i) + \tilde{R} \leq \Omega(P, B_c, C_c), \quad (1.9)$$

it follows that

$$J(A_c, B_c, C_c) \leq \text{tr } P \tilde{V}. \quad (1.10)$$

In (1.9) and (1.10) \tilde{R} and \tilde{V} denote closed-loop weighting and disturbance intensity matrices. The idea of bounding the performance by means of a Lyapunov function is the basis for guaranteed cost control ([14,17]).*

* It is also interesting to note that in Hamilton-Jacobi-Bellman sufficiency theory the performance functional is expressed in terms of a value function which also serves as a Lyapunov function for the closed-loop system. These connections will be explored in a future paper.

5. Construction of the Lyapunov Function. Thus far the Lyapunov function has only been abstractly characterized by means of (1.3) and (1.4). To obtain a useful design theory Ω is now given a concrete form. Specifically, to satisfy (1.9) it is assumed that

$$\Omega(\mathcal{P}, B_c, C_c) = \sum_{i=1}^p \Lambda_i(\mathcal{P}, B_c, C_c) + \tilde{R}, \quad (1.11)$$

where, for each i , the Λ_i are chosen such that

$$\sigma_i(\tilde{A}_i^T \mathcal{P} + \mathcal{P} \tilde{A}_i) \leq \Lambda_i(\mathcal{P}, B_c, C_c), \quad |\sigma_i| \leq \delta_i. \quad (1.12)$$

Note that (1.12) implies that (1.4) holds with Ω given by (1.11). Since \tilde{A}_i depends upon B_c and C_c , the bound Λ_i will be constructed to be *gain-invariant*, that is, so that (1.12) holds for *all* B_c and C_c . Thus, no difficulty will arise from the fact that the controller gains are yet to be determined by optimality considerations.

It should be noted that the bounding in (1.12) is defined in the sense of the cone of nonnegative-definite matrices. Since this is only a partial ordering and not a total ordering, a least upper bound (i.e., a "sharpest" bound) does not exist in general and the conservatism of the inequality in (1.12) cannot be quantified by a scalar measure. Hence, Λ_i satisfying (1.12) is not necessarily unique and two particular choices of Λ_i are developed in this paper. Since we shall utilize first-order necessary conditions for optimality, we confine our consideration to bounds which are differentiable. The first choice of Λ_i satisfying (1.12) is given by the linear (in \mathcal{P}) function

$$\Lambda_i(\mathcal{P}, B_c, C_c) = \delta_i(\alpha_i \mathcal{P} + \alpha_i^{-1} \tilde{A}_i^T \mathcal{P} \tilde{A}_i), \quad (1.13)$$

where α_i is an arbitrary positive number. As shown in [33], the bound (1.13) can be viewed as arising from a stochastic optimal control problem with exponentially weighted cost and state-, control- and measurement-dependent white noise. The stochastic multiplicative white noise model serves only as an *interpretation*, however, and need not be viewed as having physical significance. A similar bound is utilized in [28].

The second choice for Λ_i satisfying (1.12) is given by the quadratic (in \mathcal{P}) function

$$\Lambda_i(\mathcal{P}, B_c, C_c) = \delta_i(\tilde{E}_i^T \tilde{E}_i + \mathcal{P} \tilde{D}_i \tilde{D}_i^T \mathcal{P}), \quad (1.14)$$

where \tilde{D}_i, \tilde{E}_i denote an arbitrary factorization of \tilde{A}_i of the form

$$\tilde{A}_i = \tilde{D}_i \tilde{E}_i. \quad (1.15)$$

The bound (1.14) was utilized in [26] for full-state feedback with rank-1 uncertainties. Note that by utilizing congruence transformations it can be seen that both bounds (1.13) and (1.14) are basis independent. That is, replacing \tilde{A}_i by $\tilde{S}\tilde{A}_i\tilde{S}^{-1}$ leads to replacing P by $\tilde{S}^{-T}P\tilde{S}^{-1}$.

6. The Auxiliary Minimization Problem. The next step in our development for robust performance is the following. Inasmuch as the performance of a robustly stabilizing controller is bounded via (1.10) over the given range of parameter variations, it is desirable to minimize the upper bound

$$J(P, A_c, B_c, C_c) \triangleq \text{tr } P\tilde{V} \quad (1.16)$$

subject to the constraint (1.3). This is referred to as the Auxiliary Minimization Problem. For a given choice (1.13) or (1.14) of Λ_i for each i , a solution of the Auxiliary Minimization Problem provides a controller whose steady-state performance is guaranteed to remain below the bound (1.16) over the range of parameter variations, hence guaranteeing robust performance. Since the Auxiliary Minimization Problem is a smooth mathematical programming problem, a minimum always exists on compact sets. To actually characterize extremals of the Auxiliary Minimization Problem we proceed by deriving first-order necessary conditions. Because these necessary conditions are derived for the Auxiliary Minimization Problem, they effectively serve as sufficient conditions for robustness in the original problem.

It should be noted that the guaranteed cost control approach developed in [14] does not permit this line of development since Λ_i is given by

$$\Lambda_i(P, B_c, C_c) = \delta_i | \tilde{A}_i^T P + P \tilde{A}_i |, \quad (1.17)$$

where $| \cdot |$ denotes the matrix obtained by replacing each eigenvalue by its absolute value. Since this bound is not differentiable with respect to the controller gains, first-order necessary conditions cannot be used.

7. The Optimality Conditions: Full-Order Case. For the full-order case, i.e., when the order of the controller is equal to the order of the plant, the first-order necessary conditions can be derived in a form which is a direct generalization of the pair of separated Riccati equations of LQG theory. Specifically, the necessary conditions comprise a coupled system of four algebraic matrix equations including a pair of modified Riccati equations and a pair of Lyapunov equations. For plant models involving multiplicative white noise these equations have been studied in [34-36]. This form of the equations thus essentially corresponds to choosing bound (1.13).

8. The Optimality Conditions: Reduced-Order Case. For design flexibility we also consider controllers of arbitrary reduced dimension. For the linear-quadratic problem without parameter uncertainty, the formulation of the necessary conditions given in [29] provides a generalization of LQG theory. Specifically, the optimal gains are characterized by a system of algebraic matrix equations consisting of a pair of modified Riccati equations and a pair of modified Lyapunov equations coupled by an oblique projection. When the order of the controller is equal to the order of the plant, the projection becomes the identity and the standard LQG result is recovered.

The outcome of the above development is a set of algebraic matrix equations which correspond to the necessary conditions for the Auxiliary Minimization Problem and hence to sufficient conditions for robust stability and performance. These necessary conditions characterize full- or reduced-order controllers with either choice of bounds (1.13) and (1.14) for each uncertain parameter. For control-system design, these equations can be used as follows. If a solution to the necessary conditions is obtained computationally and if certain definiteness conditions hold, *then* the explicitly synthesized controller 1) solves the Robust Stability Problem and 2) is guaranteed to provide robust performance bounded by $\text{tr } P\tilde{V}$ over the stipulated uncertainty range.

The applicability of these results is, of course, limited to plants which are nominally stabilizable via controllers of the given order. Indeed, in this case it has been shown in [37] via topological degree theory that the optimality conditions for the case $\delta_i = 0$, $i = 1, \dots, p$, possess at least one stabilizing solution. For the parameter uncertainty problem, i.e., $\delta_i > 0$, it follows from continuity properties that a solution also exists for sufficiently small δ_i . The *actual* range of uncertainty which can be stabilized and the tightness of the performance bound depend upon the *conservatism* of our bounds. As will be seen from a numerical example, our bounds are not generally sharp. This is not unexpected, however, due to both the sense of the partial ordering employed in (1.12) and the fact that our choice of gain-invariant bounds permits a one-step, *non-iterative* synthesis (rather than analysis) procedure. It should be noted that necessary and sufficient conditions for robust analysis of a block-structured class of uncertainties are obtainable using Doyle's μ -function ([6]). This block structure, however, does not appear to include either the linear uncertainty model (1.1) or the matched uncertainty model of [21] as special cases.

In the present paper we present results of an illustrative numerical study for a well-known example of Doyle used in [2] to demonstrate the lack of gain margin for LQG controllers. This type of uncertainty is a special case of (1.1) obtained by taking $p = m$ and defining B_i to be the

matrix whose i th column is the same as the i th column of B and zero otherwise. To obtain full-order, robustified controllers exhibiting performance/robustness tradeoffs, we utilize bound (1.13) for several values of δ_i . To obtain these numerical results we utilized a straightforward iterative algorithm which requires only an LQG-type software package. The homotopy algorithm of [37] with appropriate extensions can also be used. Further descriptions of related algorithms and numerical results can be found in [38-40].

The development herein is self-contained with the exception that the detailed derivation of the optimality conditions has been omitted. In specialized cases the derivation has been given previously. For the case of bound (1.13) only, a derivation using Kronecker products appears in [36]. Also, a derivation without parameter uncertainties has been given in [29] using Lagrange multipliers. Overall, the derivation involves considerable matrix manipulation. Since the detailed derivation does not appear to warrant the required space, we give an outline of the proof to assist the sufficiently motivated reader in reconstructing the details.

2. Notation and Definitions

Note: All matrices have real entries

| | |
|---|---|
| $\mathbb{R}, \mathbb{R}^{r \times s}, \mathbb{R}^r, \mathbb{E}$ | real numbers, $r \times s$ real matrices, $\mathbb{R}^{r \times 1}$, expectation |
| $\ \cdot\ $ | Euclidean vector norm |
| $I_r, 0_{r \times s}, 0_r$ | $r \times r$ identity matrix, $r \times s$ zero matrix, $0_{r \times r}$ |
| $(\cdot)^T, (\cdot)^{-1}, (\cdot)^{-T}$ | transpose, inverse, inverse transpose |
| tr | trace |
| \oplus, \otimes | Kronecker sum, Kronecker product ([41]) |
| \mathbb{S}^r | $r \times r$ symmetric matrices |
| \mathbb{N}^r | $r \times r$ symmetric nonnegative-definite matrices |
| \mathbb{P}^r | $r \times r$ symmetric positive-definite matrices |
| $Z_1 \geq Z_2$ | $Z_1 - Z_2 \in \mathbb{N}^r, \quad Z_1, Z_2 \in \mathbb{S}^r$ |
| $Z_1 > Z_2$ | $Z_1 - Z_2 \in \mathbb{P}^r, \quad Z_1, Z_2 \in \mathbb{S}^r$ |
| asymptotically stable matrix | matrix with eigenvalues in open left half plane |
| n, m, ℓ, p, n_i, m_i | positive integers, $i \in \{1, \dots, p\}$ |
| \tilde{n}, \tilde{n}_i | $n + n_c, n_i + m_i, \quad i \in \{1, \dots, p\}$ |
| x, u, y, x_c | n, m, ℓ, n_c -dimensional vectors |
| $A, A_i; B, B_i; C, C_i$ | $n \times n$ matrices; $n \times m$ matrices; $\ell \times n$ matrices; $i \in \{1, \dots, p\}$ |

| | |
|---------------------------------|---|
| A_c, B_c, C_c | $n_c \times n_c, n_c \times \ell, m \times n_c$ matrices |
| \bar{A}, \bar{A}_i | $\begin{bmatrix} A & BC_c \\ B_c C & A_c \end{bmatrix} \quad \begin{bmatrix} A_i & B_i C_c \\ B_i C_i & 0 \end{bmatrix}, \quad i \in \{1, \dots, p\}$ |
| δ_i | positive number, $i \in \{1, \dots, p\}$ |
| Δ | $[-\delta_1, \delta_1] \times \dots \times [-\delta_p, \delta_p]$ |
| σ_i | real number, $i \in \{1, \dots, p\}$ |
| σ | $(\sigma_1, \dots, \sigma_p)$ |
| $\hat{\sigma}_i(\cdot)$ | Lebesgue measurable function on $[0, \infty)$, $i \in \{1, \dots, p\}$ |
| $\hat{\sigma}(\cdot)$ | $(\hat{\sigma}_1(\cdot), \dots, \hat{\sigma}_p(\cdot))$ |
| $L_\infty([0, \infty), \Delta)$ | Lebesgue measurable functions on $[0, \infty)$ with values in Δ |
| α_i | positive number, $i \in \{1, \dots, p\}$ |
| D_i, E_i, H_i, K_i | $n \times n_i, n_i \times n, n \times m_i, m_i \times m$ matrices, $i \in \{1, \dots, p\}$ |
| \tilde{D}_i, \tilde{E}_i | $\tilde{n} \times \tilde{n}_i, \tilde{n}_i \times \tilde{n}$ matrices, $i \in \{1, \dots, p\}$ |
| Σ', Σ'' | see Section 6 |
| R_1 | state weighting matrix in \mathbb{R}^n |
| R_2 | control weighting matrix in \mathbb{R}^m |
| R_{12} | $n \times m$ cross weighting matrix such that $R_1 - R_{12} R_2^{-1} R_{12}^T \geq 0$ |
| \tilde{R} | $\begin{bmatrix} R_1 & R_{12} C_c \\ C_c^T R_{12}^T & C_c^T R_2 C_c \end{bmatrix}$ |
| $w_1(\cdot)$ | n -dimensional white noise |
| $w_2(\cdot)$ | ℓ -dimensional white noise |
| V_1 | intensity of $w_1(\cdot)$ in \mathbb{R}^n |
| V_2 | intensity of $w_2(\cdot)$ in \mathbb{R}^ℓ |
| V_{12} | $n \times \ell$ cross intensity of $w_1(\cdot), w_2(\cdot)$ |
| \tilde{V} | $\begin{bmatrix} V_1 & V_{12} B_c^T \\ B_c V_{12}^T & B_c V_2 B_c^T \end{bmatrix}$ |

3. Robust Stability and Robust Performance Problems

In this section we state the Robust Stability Problem and Robust Performance Problem along with related notation for later use.

Robust Stability Problem. For fixed $n_c \leq n$, determine $(A_c, B_c, C_c) \in \mathbb{R}^{n_c \times n_c} \times \mathbb{R}^{n_c \times \ell} \times \mathbb{R}^{m \times n_c}$ such that the closed-loop system consisting of the n th-order controlled plant

$$\dot{x}(t) = \left(A + \sum_{i=1}^p \hat{\sigma}_i(t) A_i \right) x(t) + \left(B + \sum_{i=1}^p \hat{\sigma}_i(t) B_i \right) u(t), \quad \text{a.a. } t \in [0, \infty), \quad (3.1)$$

measurements

$$y(t) = \left(C + \sum_{i=1}^p \hat{\sigma}_i(t) C_i \right) x(t), \quad (3.2)$$

and n_c th-order dynamic compensator

$$\dot{x}_c(t) = A_c x_c(t) + B_c y(t), \quad (3.3)$$

$$u(t) = C_c x_c(t), \quad (3.4)$$

is asymptotically stable* for all $\hat{\sigma}(\cdot) \in L_\infty([0, \infty), \Delta)$.

Robust Performance Problem. For fixed $n_c \leq n$, determine $(A_c, B_c, C_c) \in \mathbb{R}^{n_c \times n_c} \times \mathbb{R}^{n_c \times \ell} \times \mathbb{R}^{m \times n_c}$ such that, for the closed-loop system consisting of the n th-order controlled and disturbed plant

$$\dot{x}(t) = \left(A + \sum_{i=1}^p \hat{\sigma}_i(t) A_i \right) x(t) + \left(B + \sum_{i=1}^p \hat{\sigma}_i(t) B_i \right) u(t) + w_1(t), \quad \text{a.a. } t \in [0, \infty), \quad (3.5)$$

noisy measurements

$$y(t) = \left(C + \sum_{i=1}^p \hat{\sigma}_i(t) C_i \right) x(t) + w_2(t), \quad (3.6)$$

and n_c th-order dynamic compensator (3.3), (3.4), the performance criterion

$$J(A_c, B_c, C_c) \triangleq \sup_{\hat{\sigma}(\cdot) \in L_\infty([0, \infty), \Delta)} \limsup_{t \rightarrow \infty} \mathbb{E} [x^T(t) R_1 x(t) + 2x^T(t) R_{12} u(t) + u^T(t) R_2 u(t)] \quad (3.7)$$

is minimized.

For each controller (A_c, B_c, C_c) and parameter variation $\hat{\sigma}(\cdot) \in L_\infty([0, \infty), \Delta)$ the undisturbed closed-loop system (3.1)–(3.4) is given by

$$\dot{\tilde{x}}(t) = \left(\tilde{A} + \sum_{i=1}^p \hat{\sigma}_i(t) \tilde{A}_i \right) \tilde{x}(t), \quad \text{a.a. } t \in [0, \infty), \quad (3.8)$$

while the disturbed closed-loop system (3.3)–(3.6) is

$$\dot{\tilde{x}}(t) = \left(\tilde{A} + \sum_{i=1}^p \hat{\sigma}_i(t) \tilde{A}_i \right) \tilde{x}(t) + \tilde{w}(t), \quad \text{a.a. } t \in [0, \infty). \quad (3.9)$$

* Asymptotic stability for a nonautonomous system is defined in the standard way. See, e.g., [42].

Also (see, e.g., [43], p. 194), let $\tilde{\Phi} : [0, \infty) \rightarrow \mathbb{R}^{\tilde{n} \times \tilde{n}}$ be the unique absolutely continuous solution to

$$\dot{\tilde{\Phi}}(t) = \left(\tilde{A} + \sum_{i=1}^p \hat{\sigma}_i(t) \tilde{A}_i \right) \tilde{\Phi}(t), \quad \text{a.a. } t \in [0, \infty), \quad (3.10)$$

$$\tilde{\Phi}(0) = I_{\tilde{n}}, \quad (3.11)$$

and recall that $\tilde{\Phi}^{-1}(\cdot)$ satisfies

$$\frac{d}{dt} \tilde{\Phi}^{-1}(t) = -\tilde{\Phi}^{-1}(t) \left(\tilde{A} + \sum_{i=1}^p \hat{\sigma}_i(t) \tilde{A}_i \right), \quad \text{a.a. } t \in [0, \infty). \quad (3.12)$$

4. Sufficient Conditions for Robust Stability and Performance

For robust stability we characterize quadratic Lyapunov functions for the closed-loop system.

Theorem 4.1. Let $\Omega : \mathbb{P}^{\tilde{n}} \times \mathbb{R}^{n_o \times \ell} \times \mathbb{R}^{m \times n_o} \rightarrow \mathbb{S}^{\tilde{n}}$ satisfy

$$\sum_{i=1}^p \sigma_i(\tilde{A}_i^T \rho + \rho \tilde{A}_i) < \Omega(\rho, B_c, C_c), \quad \sigma \in \Delta, \quad (\rho, B_c, C_c) \in \mathbb{P}^{\tilde{n}} \times \mathbb{R}^{n_o \times \ell} \times \mathbb{R}^{m \times n_o}. \quad (4.1)$$

If, for some $(A_c, B_c, C_c) \in \mathbb{R}^{n_o \times n_o} \times \mathbb{R}^{n_o \times \ell} \times \mathbb{R}^{m \times n_o}$, there exists $\rho \in \mathbb{P}^{\tilde{n}}$ satisfying

$$0 = \tilde{A}^T \rho + \rho \tilde{A} + \Omega(\rho, B_c, C_c), \quad (4.2)$$

then (A_c, B_c, C_c) solves the Robust Stability Problem.

Proof. Define the Lyapunov function

$$V(\tilde{x}) \triangleq \tilde{x}^T \rho \tilde{x}, \quad \tilde{x} \in \mathbb{R}^{\tilde{n}}.$$

For $t \in [0, \infty)$ and $\tilde{x}(t)$ satisfying (3.8), it follows from (4.2) that

$$\begin{aligned} \dot{V}(\tilde{x}(t)) &= \dot{\tilde{x}}^T(t) \rho \tilde{x}(t) + \tilde{x}^T(t) \rho \dot{\tilde{x}}(t) \\ &= \tilde{x}^T(t) \left[\left(\tilde{A} + \sum_{i=1}^p \hat{\sigma}_i(t) \tilde{A}_i \right)^T \rho + \rho \left(\tilde{A} + \sum_{i=1}^p \hat{\sigma}_i(t) \tilde{A}_i \right) \right] \tilde{x}(t) \\ &= \tilde{x}^T(t) \left[\sum_{i=1}^p \hat{\sigma}_i(t) (\tilde{A}_i^T \rho + \rho \tilde{A}_i) - \Omega(\rho, B_c, C_c) \right] \tilde{x}(t). \end{aligned}$$

Since $\hat{\sigma}(t) \in \Delta$, $t \in [0, \infty)$, it follows from (4.1) that there exists $\gamma > 0$ such that $\dot{V}(\tilde{x}(t)) \leq -\gamma \|\tilde{x}(t)\|^2$, $t \in [0, \infty)$. \square

Remark 4.1. If (A_c, B_c, C_c) solves the Robust Stability Problem then

$$\lim_{t \rightarrow \infty} \tilde{\Phi}(t) = 0, \quad \partial(\cdot) \in L_\infty([0, \infty), \Delta). \quad (4.3)$$

Remark 4.2. As will be seen the bound (4.1) will be guaranteed for all P, B_c, C_c by suitable construction of the function Ω . In addition, the existence of a solution P to (4.2) need not be verified in practice. Rather, (4.2) is a result of numerically solving the necessary conditions for the Auxiliary Minimization Problem given by Theorem 6.1.

For the Robust Performance Problem the cost can be expressed in terms of the closed-loop second-moment matrix.

Proposition 4.1. For $(A_c, B_c, C_c) \in \mathbb{R}^{n_0 \times n_0} \times \mathbb{R}^{n_0 \times \ell} \times \mathbb{R}^{m \times n_0}$ and $\partial(\cdot) \in L_\infty([0, \infty), \Delta)$ the second-moment matrix

$$\tilde{Q}(t) \triangleq \mathbb{E}[\tilde{z}(t)\tilde{z}^T(t)], \quad t \in [0, \infty), \quad (4.4)$$

satisfies

$$\dot{\tilde{Q}}(t) = \left(\tilde{A} + \sum_{i=1}^p \partial_i(t) \tilde{A}_i \right) \tilde{Q}(t) + \tilde{Q}(t) \left(\tilde{A} + \sum_{i=1}^p \partial_i(t) \tilde{A}_i \right)^T + \tilde{V}, \quad \text{a.a. } t \in [0, \infty), \quad (4.5)$$

or, equivalently,

$$\tilde{Q}(t) = \tilde{\Phi}(t) \tilde{Q}(0) \tilde{\Phi}^T(t) + \int_0^t \tilde{\Phi}(t) \tilde{\Phi}^{-1}(s) \tilde{V} \tilde{\Phi}^{-T}(s) \tilde{\Phi}^T(t) ds, \quad t \in [0, \infty). \quad (4.6)$$

Furthermore,

$$J(A_c, B_c, C_c) = \sup_{\partial(\cdot) \in L_\infty([0, \infty), \Delta)} \limsup_{t \rightarrow \infty} \text{tr } \tilde{Q}(t) \tilde{R}, \quad (4.7)$$

or, equivalently,

$$J(A_c, B_c, C_c) \triangleq \sup_{\partial(\cdot) \in L_\infty([0, \infty), \Delta)} \limsup_{t \rightarrow \infty} \text{tr} \left[\tilde{\Phi}(t) \tilde{Q}(0) \tilde{\Phi}^T(t) \tilde{R} + \int_0^t \tilde{\Phi}(t) \tilde{\Phi}^{-1}(s) \tilde{V} \tilde{\Phi}^{-T}(s) \tilde{\Phi}^T(t) ds \tilde{R} \right]. \quad (4.8)$$

Proof. The second-moment equation (4.5) is a direct consequence of the Ito differential rule (see [44], p. 142) while (4.6) follows by direct verification. Finally, (4.7) is immediate. \square

We now proceed to obtain an upper bound for J in terms of the matrix P . The following lemma is required.

Lemma 4.1. Let $\Omega : \mathbb{P}^n \times \mathbb{R}^{n \times \ell} \times \mathbb{R}^{m \times n} \rightarrow \mathbb{S}^n$ and $(A_c, B_c, C_c) \in \mathbb{R}^{n \times n} \times \mathbb{R}^{n \times \ell} \times \mathbb{R}^{m \times n}$ be given. Then $P \in \mathbb{P}^n$ satisfies (4.2) if and only if P satisfies

$$\begin{aligned} P &= \tilde{\Phi}^T(t) P \tilde{\Phi}(t) \\ &+ \int_0^t \tilde{\Phi}^T(t) \tilde{\Phi}^{-T}(s) \left[\Omega(P, B_c, C_c) - \sum_{i=1}^p \hat{\sigma}_i(t) (\tilde{A}_i^T P + P \tilde{A}_i) \right] \tilde{\Phi}^{-1}(s) \tilde{\Phi}(t) ds, \end{aligned} \quad (4.9)$$

$$\hat{\sigma}(\cdot) \in L_\infty([0, \infty), \Delta), \quad t \in [0, \infty).$$

Proof. Suppose P satisfies (4.2). Then, for $t \in [0, \infty)$,

$$\begin{aligned} 0 &= \tilde{\Phi}^{-T}(t) \left(\tilde{A} + \sum_{i=1}^p \hat{\sigma}_i(t) \tilde{A}_i \right)^T P \tilde{\Phi}^{-1}(t) + \tilde{\Phi}^{-T}(t) P \left(\tilde{A} + \sum_{i=1}^p \hat{\sigma}_i(t) \tilde{A}_i \right) \tilde{\Phi}^{-1}(t) \\ &+ \tilde{\Phi}^{-T}(t) \left[\Omega(P, B_c, C_c) - \sum_{i=1}^p \hat{\sigma}_i(t) (\tilde{A}_i^T P + P \tilde{A}_i) \right] \tilde{\Phi}^{-1}(t) \\ &= -\frac{d}{dt} \left[\tilde{\Phi}^{-T}(t) P \tilde{\Phi}^{-1}(t) \right] + \tilde{\Phi}^{-T}(t) \left[\Omega(P, B_c, C_c) - \sum_{i=1}^p \hat{\sigma}_i(t) (\tilde{A}_i^T P + P \tilde{A}_i) \right] \tilde{\Phi}^{-1}(t), \end{aligned}$$

which yields

$$0 = -\tilde{\Phi}^{-T}(t) P \tilde{\Phi}^{-1}(t) + P + \int_0^t \tilde{\Phi}^{-T}(s) \left[\Omega(P, B_c, C_c) - \sum_{i=1}^p \hat{\sigma}_i(s) (\tilde{A}_i^T P + P \tilde{A}_i) \right] \tilde{\Phi}^{-1}(s) ds.$$

Thus (4.9) is satisfied. Conversely, suppose P satisfies (4.9). Differentiating with respect to t using Leibniz' rule yields

$$\begin{aligned} 0 &= \left(\tilde{A} + \sum_{i=1}^p \hat{\sigma}_i(t) \tilde{A}_i \right)^T \tilde{\Phi}^T(t) P \tilde{\Phi}(t) + \tilde{\Phi}^T(t) P \tilde{\Phi}(t) \left(\tilde{A} + \sum_{i=1}^p \hat{\sigma}_i(t) \tilde{A}_i \right) \\ &+ \left(\tilde{A} + \sum_{i=1}^p \hat{\sigma}_i(t) \tilde{A}_i \right)^T \int_0^t \tilde{\Phi}^T(t) \tilde{\Phi}^{-T}(s) \left[\Omega(P, B_c, C_c) - \sum_{i=1}^p \hat{\sigma}_i(s) (\tilde{A}_i^T P + P \tilde{A}_i) \right] \tilde{\Phi}^{-1}(s) \tilde{\Phi}(t) ds \\ &+ \int_0^t \tilde{\Phi}^T(t) \tilde{\Phi}^{-T}(s) \left[\Omega(P, B_c, C_c) - \sum_{i=1}^p \hat{\sigma}_i(s) (\tilde{A}_i^T P + P \tilde{A}_i) \right] \tilde{\Phi}^{-1}(s) \tilde{\Phi}(t) ds \left(\tilde{A} + \sum_{i=1}^p \hat{\sigma}_i(t) \tilde{A}_i \right) \\ &+ \Omega(P, B_c, C_c) - \sum_{i=1}^p \hat{\sigma}_i(t) (\tilde{A}_i^T P + P \tilde{A}_i) \\ &= \left(\tilde{A} + \sum_{i=1}^p \hat{\sigma}_i(t) \tilde{A}_i \right)^T P + P \left(\tilde{A} + \sum_{i=1}^p \hat{\sigma}_i(t) \tilde{A}_i \right) + \Omega(P, B_c, C_c) - \sum_{i=1}^p \hat{\sigma}_i(t) (\tilde{A}_i^T P + P \tilde{A}_i) \\ &= \tilde{A}^T P + P \tilde{A} + \Omega(P, B_c, C_c). \end{aligned}$$

Hence (4.2) is satisfied. \square

Remark 4.3. Note the identity

$$\text{tr} \int_0^t \tilde{\Phi}(t) \tilde{\Phi}^{-1}(s) \tilde{V} \tilde{\Phi}^{-T}(s) \tilde{\Phi}^T(t) ds \tilde{R} = \text{tr} \int_0^t \tilde{\Phi}^T(t) \tilde{\Phi}^{-T}(s) \tilde{R} \tilde{\Phi}^{-1}(s) \tilde{\Phi}(t) ds \tilde{V}, \quad (4.10)$$

$$(A_c, B_c, C_c) \in \mathbb{R}^{n_c \times n_c} \times \mathbb{R}^{n_c \times \ell} \times \mathbb{R}^{m \times n_c}, \quad \hat{\sigma}(\cdot) \in L_\infty([0, \infty), \Delta), \quad t \in [0, \infty).$$

We are now in a position to bound the cost J by means of the matrix P .

Theorem 4.2. Let $\Omega : \mathbb{IP}^{\tilde{n}} \times \mathbb{R}^{n_c \times \ell} \times \mathbb{R}^{m \times n_c} \rightarrow \mathbb{S}^{\tilde{n}}$ satisfy (4.1) and

$$\sum_{i=1}^p \sigma_i(\tilde{A}_i^T P + P \tilde{A}_i) + \tilde{R} \leq \Omega(P, B_c, C_c), \quad \sigma \in \Delta, \quad (P, B_c, C_c) \in \mathbb{IP}^{\tilde{n}} \times \mathbb{R}^{n_c \times \ell} \times \mathbb{R}^{m \times n_c}. \quad (4.11)$$

If, for some $(A_c, B_c, C_c) \in \mathbb{R}^{n_c \times n_c} \times \mathbb{R}^{n_c \times \ell} \times \mathbb{R}^{m \times n_c}$, there exists $P \in \mathbb{IP}^{\tilde{n}}$ satisfying (4.2), then

$$J(A_c, B_c, C_c) \leq \text{tr } P \tilde{V}. \quad (4.12)$$

Proof. From (4.8)–(4.10) and (4.3) it follows that

$$\begin{aligned} J(A_c, B_c, C_c) &= \sup_{\hat{\sigma}(\cdot) \in L_\infty([0, \infty), \Delta)} \limsup_{t \rightarrow \infty} \text{tr} \left\{ \tilde{\Phi}(t) \tilde{Q}(0) \tilde{\Phi}^T(t) \tilde{R} + P \tilde{V} - \tilde{\Phi}^T(t) P \tilde{\Phi}(t) \tilde{V} \right. \\ &\quad \left. - \int_0^t \tilde{\Phi}^T(s) \tilde{\Phi}^{-T}(s) [\Omega(P, B_c, C_c) - \tilde{R} - \sum_{i=1}^p \hat{\sigma}_i(s) (\tilde{A}_i^T P + P \tilde{A}_i)] \tilde{\Phi}^{-1}(s) \tilde{\Phi}(s) ds \tilde{V} \right\} \\ &\leq \sup_{\hat{\sigma}(\cdot) \in L_\infty([0, \infty), \Delta)} \limsup_{t \rightarrow \infty} \text{tr} [\tilde{\Phi}(t) \tilde{Q}(0) \tilde{\Phi}^T(t) \tilde{R} + P \tilde{V}] \\ &= \text{tr } P \tilde{V}. \quad \square \end{aligned}$$

Remark 4.4. Note that since $\tilde{R} \geq 0$, (4.11) implies

$$\sum_{i=1}^p \sigma_i(\tilde{A}_i^T P + P \tilde{A}_i) \leq \Omega(P, B_c, C_c), \quad \sigma \in \Delta, \quad (4.13)$$

which is a weak form of (4.1). If $\tilde{R} > 0$ then (4.11) implies (4.1). This implication is not surprising since (4.11) implies robust performance while (4.1) implies robust stability.

5. Choice of Bounds

To satisfy (4.11) $\Omega(\cdot, \cdot, \cdot)$ is chosen to be of the form

$$\Omega(P, B_c, C_c) = \sum_{i=1}^p \Lambda_i(P, B_c, C_c) + \tilde{R}, \quad (5.1)$$

where, for each $i = 1, \dots, p$, $\Lambda_i : \mathbb{IP}^{\tilde{n}} \times \mathbb{R}^{n_c \times \ell} \times \mathbb{R}^{m \times n_c} \rightarrow \mathbb{S}^{\tilde{n}}$ satisfies

$$\sigma_i(\tilde{A}_i^T P + P \tilde{A}_i) \leq \Lambda_i(P, B_c, C_c), \quad \sigma_i \in [-\delta_i, \delta_i], \quad (P, B_c, C_c) \in \mathbb{IP}^{\tilde{n}} \times \mathbb{R}^{n_c \times \ell} \times \mathbb{R}^{m \times n_c}. \quad (5.2)$$

Two distinct choices for the bound Λ_i are considered. As pointed out in Section 1, the first choice corresponds to a right shift/multiplicative white noise model ([33]), while the second bound generalizes results found in [26].

Proposition 5.1. For all $\alpha_i > 0$ the function

$$\Lambda_i(\mathcal{P}, B_c, C_c) = \delta_i(\alpha_i \mathcal{P} + \alpha_i^{-1} \tilde{A}_i^T \mathcal{P} \tilde{A}_i) \quad (5.3)$$

satisfies (5.2).

Proof. Note that

$$\begin{aligned} 0 &\leq [\sigma_i(\alpha_i/\delta_i)^{\frac{1}{2}} I_{\tilde{n}} - (\delta_i/\alpha_i)^{\frac{1}{2}} \tilde{A}_i]^T \mathcal{P} [\sigma_i(\alpha_i/\delta_i)^{\frac{1}{2}} I_{\tilde{n}} - (\delta_i/\alpha_i)^{\frac{1}{2}} \tilde{A}_i] \\ &= \sigma_i^2(\alpha_i/\delta_i) \mathcal{P} + (\delta_i/\alpha_i) \tilde{A}_i^T \mathcal{P} \tilde{A}_i - \sigma_i(\tilde{A}_i^T \mathcal{P} + \mathcal{P} \tilde{A}_i), \end{aligned}$$

which, since $\alpha_i^2 \leq \delta_i^2$, implies (5.2). \square

Proposition 5.2. For all $\tilde{D}_i \in \mathbb{R}^{\tilde{n} \times \tilde{n}_i}$ and $\tilde{E}_i \in \mathbb{R}^{\tilde{n}_i \times \tilde{n}}$ satisfying

$$\tilde{A}_i = \tilde{D}_i \tilde{E}_i, \quad (5.4)$$

the function

$$\Lambda_i(\mathcal{P}, B_c, C_c) = \delta_i(\tilde{E}_i^T \tilde{E}_i + \mathcal{P} \tilde{D}_i \tilde{D}_i^T \mathcal{P}) \quad (5.5)$$

satisfies (5.2).

Proof. Note that

$$\begin{aligned} 0 &\leq [\delta_i^{\frac{1}{2}} \tilde{E}_i - \sigma_i \delta_i^{-\frac{1}{2}} \tilde{D}_i^T \mathcal{P}]^T [\delta_i^{\frac{1}{2}} \tilde{E}_i - \sigma_i \delta_i^{-\frac{1}{2}} \tilde{D}_i^T \mathcal{P}] \\ &= \delta_i \tilde{E}_i^T \tilde{E}_i + (\sigma_i^2/\delta_i) \mathcal{P} \tilde{D}_i \tilde{D}_i^T \mathcal{P} - \sigma_i(\tilde{A}_i^T \mathcal{P} + \mathcal{P} \tilde{A}_i), \end{aligned}$$

which implies (5.2). \square

6. The Auxiliary Minimization Problem and Necessary Conditions for Optimality

To optimize robust performance while retaining robust stability, we consider the following problem for which the cost functional is given by the bound (4.12).

Auxiliary Minimization Problem. For $i = 1, \dots, p$ let Λ_i be given by either (5.3) or (5.5). Determine $(\mathcal{P}, A_c, B_c, C_c) \in \mathbb{P}^{\tilde{n}} \times \mathbb{R}^{n_o \times n_o} \times \mathbb{R}^{n_o \times \ell} \times \mathbb{R}^{m \times n_o}$ which minimizes

$$J(\mathcal{P}, A_c, B_c, C_c) \triangleq \text{tr } \mathcal{P} \tilde{V} \quad (6.1)$$

subject to

$$0 = \tilde{A}^T P + P \tilde{A} + \sum_{i=1}^p \Lambda_i(P, B_c, C_c) + \tilde{R} \quad (6.2)$$

and

$$\sum_{i=1}^p \sigma_i(\tilde{A}_i^T P + P \tilde{A}_i) < \sum_{i=1}^p \Lambda_i(P, B_c, C_c) + \tilde{R}, \quad \sigma \in \Delta. \quad (6.3)$$

Remark 6.1. Note that (6.3) enforces both (4.1) and (4.11) to guarantee robust stability and performance.

To derive first-order necessary conditions for the Auxiliary Minimization Problem, note that the constraint (6.3) defines an open set.

Proposition 6.1. The set of $(P, B_c, C_c) \in \mathbb{P}^{\tilde{n}} \times \mathbb{R}^{n_c \times \ell} \times \mathbb{R}^{m \times n_c}$ satisfying (6.3) is open.

Proof. Since $\Lambda_i(\cdot, \cdot, \cdot)$ is continuous it can be shown that the function

$$f(P, B_c, C_c) \triangleq \min_{\sigma \in \Delta} \lambda_{\min} \left\{ \sum_{i=1}^p \Lambda_i(P, B_c, C_c) + \tilde{R} - \sum_{i=1}^p \sigma_i(\tilde{A}_i^T P + P \tilde{A}_i) \right\}$$

is also continuous. Since (6.3) is equivalent to $0 < f(P, B_c, C_c)$, the result is immediate. \square

To obtain explicit feedback gain expressions we shall require two additional technical assumptions. If bound (5.3) is chosen for a given $i \in \{1, \dots, p\}$ we require

$$B_i \neq 0 \implies C_i = 0, \quad (6.4)$$

i.e., B_i and C_i are not simultaneously nonzero. Of course, both B_i and C_i may be zero. Assumption (6.4) implies that parameter uncertainties in B and C must be modeled as uncorrelated. Correlation between uncertainties in A and B or A and C is, of course, permitted. Furthermore, if bound (5.5) is chosen for a given $i \in \{1, \dots, p\}$ we require

$$C_i = 0. \quad (6.5)$$

When utilizing bound (5.3) the positive constant α_i shall be considered fixed but arbitrary. Furthermore, for bound (5.5), let $D_i \in \mathbb{R}^{n \times n_i}$, $E_i \in \mathbb{R}^{n_i \times n}$, $H_i \in \mathbb{R}^{n \times m_i}$ and $K_i \in \mathbb{R}^{m_i \times m}$ satisfy

$$A_i = D_i E_i, \quad B_i = H_i K_i, \quad (6.6)$$

and define \tilde{D}_i, \tilde{E}_i satisfying (5.4) by

$$\tilde{D}_i \triangleq \begin{bmatrix} D_i & H_i \\ 0_{n_o \times n_i} & 0_{n_o \times m_i} \end{bmatrix}, \quad \tilde{E}_i \triangleq \begin{bmatrix} E_i & 0_{n_i \times n_o} \\ 0_{m_i \times n} & K_i C_c \end{bmatrix}. \quad (6.7)$$

In addition to the open set defined by (6.3), the derivation of the necessary conditions requires that (P, A_c, B_c, C_c) be further restricted so that

$$\begin{aligned} (\tilde{A} + \frac{1}{2} \sum' \delta_i \alpha_i I_{\tilde{n}} + \sum'' \delta_i \tilde{D}_i \tilde{D}_i^T P) \oplus (\tilde{A} + \frac{1}{2} \sum' \delta_i \alpha_i I_{\tilde{n}} + \sum'' \delta_i \tilde{D}_i \tilde{D}_i^T P) \\ + \sum' (\delta_i / \alpha_i) \tilde{A}_i \otimes \tilde{A}_i \quad \text{is asymptotically stable} \end{aligned} \quad (6.8)$$

and

$$(A_c, B_c, C_c) \text{ is controllable and observable.} \quad (6.9)$$

In (6.8) the notation \sum' and \sum'' denotes summation over indices for which bounds (5.3) and (5.5), respectively, have been chosen. Note that (6.8) and (6.9) play no role in the Auxiliary Minimization Problem and thus need not be verified for robust stability or robust performance.

For arbitrary $Q, P, \hat{Q}, \hat{P} \in \mathbb{R}^{n \times n}$ define the following notation:

$$\begin{aligned} R_{2a} &\triangleq R_2 + \sum' (\delta_i / \alpha_i) B_i^T (P + \hat{P}) B_i + \sum'' \delta_i K_i^T K_i, & V_{2a} &\triangleq V_2 + \sum' (\delta_i / \alpha_i) C_i (Q + \hat{Q}) C_i^T, \\ P_a &\triangleq B^T P + R_{12}^T + \sum' (\delta_i / \alpha_i) B_i^T (P + \hat{P}) A_i, & Q_a &\triangleq Q C^T + V_{12} + \sum' (\delta_i / \alpha_i) A_i (Q + \hat{Q}) C_i^T, \\ D &\triangleq \sum'' \delta_i (D_i D_i^T + H_i H_i^T), & E &\triangleq \sum'' \delta_i E_i^T E_i, \\ \hat{A} &\triangleq A + \frac{1}{2} \sum' \delta_i \alpha_i I_n, & \hat{A}_P &\triangleq \hat{A} - B R_{2a}^{-1} P_a, & \hat{A}_Q &\triangleq \hat{A} - Q_a V_{2a}^{-1} C. \end{aligned}$$

The following lemma will be needed.

Lemma 6.1. If $\hat{Q}, \hat{P} \in \mathbb{R}^{n \times n}$ and $\text{rank } \hat{Q} \hat{P} = n_o$, then there exist $G, \Gamma \in \mathbb{R}^{n_o \times n}$ and invertible $M \in \mathbb{R}^{n_o \times n_o}$ such that

$$\hat{Q} \hat{P} = G^T M \Gamma, \quad (6.10)$$

$$\Gamma G^T = I_{n_o}. \quad (6.11)$$

Furthermore, G, M and Γ are unique except for a change of basis in \mathbb{R}^{n_o} .

Proof. The result is an immediate consequence of [45], Theorem 6.2.5, p. 123. \square

Note that because of (6.11), the $n \times n$ matrix $\tau \triangleq G^T \Gamma$ is idempotent, i.e., $\tau^2 = \tau$. Since τ is not necessarily symmetric, it is an oblique projection. Also, define $\tau_{\perp} \triangleq I_n - \tau$.

Theorem 6.1. Suppose (P, A_c, B_c, C_c) solves the Auxiliary Minimization Problem subject to (6.8) and (6.9). Then there exist $P, Q, \hat{P}, \hat{Q} \in \mathbb{R}^n$ such that P, A_c, B_c, C_c are given by

$$P = \begin{bmatrix} P + \hat{P} & -\hat{P}G^T \\ -G\hat{P} & G\hat{P}G^T \end{bmatrix}, \quad (6.12)$$

$$A_c = \Gamma(A - Q_a V_{2a}^{-1} C - B R_{2a}^{-1} P_a + D P) G^T, \quad (6.13)$$

$$B_c = \Gamma Q_a V_{2a}^{-1}, \quad (6.14)$$

$$C_c = -R_{2a}^{-1} P_a G^T, \quad (6.15)$$

and such that P, Q, \hat{P}, \hat{Q} satisfy

$$0 = \hat{A}^T P + P \hat{A} + R_1 + \sum' (\delta_i / \alpha_i) [A_i^T P A_i + (A_i - Q_a V_{2a}^{-1} C_i)^T \hat{P} (A_i - Q_a V_{2a}^{-1} C_i)] \\ + E + P D P - P_a^T R_{2a}^{-1} P_a + \tau_{\perp}^T P_a^T R_{2a}^{-1} P_a \tau_{\perp}, \quad (6.16)$$

$$0 = [\hat{A} + D(P + \hat{P})] Q + Q [\hat{A} + D(P + \hat{P})]^T + V_1 + \sum' (\delta_i / \alpha_i) [A_i Q A_i^T \\ + (A_i - B_i R_{2a}^{-1} P_a) \hat{Q} (A_i - B_i R_{2a}^{-1} P_a)^T] - Q_a V_{2a}^{-1} Q_a^T + \tau_{\perp} Q_a V_{2a}^{-1} Q_a^T \tau_{\perp}^T, \quad (6.17)$$

$$0 = (\hat{A}_Q + D P)^T \hat{P} + \hat{P} (\hat{A}_Q + D P) + \hat{P} D \hat{P} + P_a^T R_{2a}^{-1} P_a - \tau_{\perp}^T P_a^T R_{2a}^{-1} P_a \tau_{\perp}, \quad (6.18)$$

$$0 = (\hat{A}_P + D P) \hat{Q} + \hat{Q} (\hat{A}_P + D P)^T + Q_a V_{2a}^{-1} Q_a^T - \tau_{\perp} Q_a V_{2a}^{-1} Q_a^T \tau_{\perp}^T, \quad (6.19)$$

$$\text{rank } \hat{Q} = \text{rank } \hat{P} = \text{rank } \hat{Q} \hat{P} = n_c. \quad (6.20)$$

Conversely, if there exist $P, Q, \hat{P}, \hat{Q} \in \mathbb{R}^n$ satisfying (6.16)–(6.20), then P given by (6.12) satisfies (6.2) or, equivalently, (4.2) with (A_c, B_c, C_c) given by (6.13)–(6.15).

Outline of Proof. As discussed in Section 1, we limit the presentation of the proof to the salient details. First note that with the choice of bounds Λ_i , (6.2) becomes

$$0 = (\tilde{A} + \frac{1}{2} \sum' \delta_i \alpha_i I_{\tilde{n}})^T P + P (\tilde{A} + \frac{1}{2} \sum' \delta_i \alpha_i I_{\tilde{n}} + \tilde{R} \\ + \sum' (\delta_i / \alpha_i) \tilde{A}_i^T P \tilde{A}_i + \sum'' \delta_i (\tilde{E}_i^T \tilde{E}_i + P \tilde{D}_i \tilde{D}_i^T P)). \quad (6.21)$$

By introducing multipliers $\lambda \in \mathbb{R}$, $\lambda \geq 0$, and $Q \in \mathbb{R}^{\tilde{n} \times \tilde{n}}$, a Lagrangian can be defined as

$$\mathcal{L}(P, A_c, B_c, C_c) \triangleq \text{tr}[\lambda P \tilde{V} + Q(\text{RHS of (6.21)})]. \quad (6.22)$$

Setting $\partial \mathcal{L} / \partial P = 0$ and utilizing (6.8) implies that $\lambda = 1$ without loss of generality, $Q \geq 0$, and Q satisfies

$$0 = (\tilde{A} + \frac{1}{2} \sum' \delta_i \alpha_i I_{\tilde{n}} + \sum'' \tilde{D}_i \tilde{D}_i^T P) Q + Q (\tilde{A} + \frac{1}{2} \sum' \delta_i \alpha_i I_{\tilde{n}} + \sum'' \tilde{D}_i \tilde{D}_i^T P)^T \\ + \sum' (\delta_i / \alpha_i) \tilde{A}_i Q \tilde{A}_i^T + \tilde{V}. \quad (6.23)$$

The remainder of the derivation is exactly parallel to the techniques utilized in [29,36]. Briefly, the principal steps are as follows:

Step 1. Compute $\partial \mathcal{L} / \partial A_c, \partial \mathcal{L} / \partial B_c$ and $\partial \mathcal{L} / \partial C_c$;

Step 2. Use (6.9) to show that the lower right $n_c \times n_c$ blocks of Q and P are positive definite;

Step 3. Use $\partial \mathcal{L} / \partial A_c = 0$ to define a projection r and new variables $P, Q, \hat{P}, \hat{Q}, G, \Gamma$;

Step 5. Partition (6.21) and (6.23) into six equations (#1, ..., #6) corresponding to the $n \times n$, $n \times n_c$ and $n_c \times n_c$ blocks of P and Q , respectively;

Step 6. Use equations #2 and #3 to solve for A_c ; show that equations #5 and #6 also yield A_c ; note that with A_c now given, equations #3 and #6 are superfluous and can be eliminated;

Step 7. Manipulate equations #1, #2, #4 and #5 to yield (6.16)–(6.19);

Step 8. Show that Steps 5–7 are reversible so that (6.16)–(6.20) are equivalent to (6.2) or, equivalently, (4.2). \square

By enforcing the strict inequalities $P > 0$ and (6.3), solutions of (6.16)–(6.20) guarantee robust stability with a robust performance bound. The following result follows from Theorem 4.1, Theorem 4.2 and the converse of Theorem 6.1.

Theorem 6.2. Suppose there exist $P, Q, \hat{P}, \hat{Q} \in \mathbb{R}^n$ satisfying (6.16)–(6.20), and suppose that (6.3) and $P > 0$ are satisfied with (P, A_c, B_c, C_c) given by (6.12)–(6.15). Then the compensator A_c, B_c, C_c given by (6.13)–(6.15) solves the Robust Stability Problem and the closed-loop performance (3.7) satisfies the bound

$$J(A_c, B_c, C_c) \leq \text{tr } P \tilde{V}. \quad (6.24)$$

The following existence result concerns the solvability of (6.16)–(6.20). Let n_u denote the dimension of the unstable subspace of the plant dynamics matrix A .

Theorem 6.3. Assume $n_c \geq n_u$, $R_1 > 0$, $V_1 > 0$, suppose the nominal plant, i.e., (3.1), (3.2) with $\delta_i = 0, i = 1, \dots, p$, is stabilizable and detectable and, in addition, is stabilizable by means of an n_c th-order strictly proper dynamic compensator (3.3), (3.4). Then there exist $\bar{\delta}_1, \dots, \bar{\delta}_p > 0$

such that if $\delta_i \in [0, \bar{\delta}_i], i = 1, \dots, p$, then (6.16)–(6.20) have a solution $P, Q, \hat{P}, \hat{Q} \in \mathbb{R}^n$ for which (A_c, B_c, C_c) given by (6.13)–(6.15) solve the robust stability problem with robust performance bound (6.24).

Proof. From Theorem 3.1 of [37] it follows that there exists a solution to (6.16)–(6.20) which stabilizes the nominal plant. By continuity there exists a neighborhood over which robust stability with performance bound (6.24) holds. \square

Theorem 6.3 is an existence result which guarantees solvability of the sufficiency conditions over a range of parameter uncertainties. The actual range of uncertainty which can be bounded and the conservatism of the performance bound are problem dependent. To this end we now consider a numerical example.

7. Illustrative Numerical Example

To demonstrate the above theory we present an illustrative numerical example. The example chosen was originally used in [2] to illustrate the lack of a guaranteed gain margin for LQG controllers. This example was also considered in [35] for a preliminary robustness study and reconsidered in [46] using μ -analysis. Define

$$\begin{aligned} n = n_u = 2, \quad m = \ell = p = 1, \\ A = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}, \quad B = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \quad C = [1 \quad 0], \\ A_1 = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}, \quad B_1 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \quad C_1 = [0 \quad 0], \\ R_1 = V_1 = \begin{bmatrix} 60 & 60 \\ 60 & 60 \end{bmatrix}, \quad R_{12} = V_{12} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \quad R_2 = V_2 = 1. \end{aligned}$$

Note that the system is open-loop unstable and becomes uncontrollable at $\sigma_1 = -1$. As can be seen using root locus, a strictly proper stabilizing controller must be of at least second order. Hence we consider (6.16)–(6.20) with $n_c = n$ and $r_1 = 0$. Furthermore, we utilize bound (5.3) and thus set $D = E = 0$. Using algorithms described in [38–40], controllers were obtained by solving (6.16)–(6.20) for $(\delta_1, \alpha_1) = (.1, 1), (.2, 2)$ and $(.4, 4)$. As stated previously, these numerical solutions also verify (4.2) with \mathcal{P} given by (6.12). Figure 1 compares the guaranteed robust stability region to the “actual” robust stability region. This robust stability region was evaluated assuming constant $\hat{\sigma}_1(\cdot)$ although the theory actually guarantees robustness with respect to time-varying uncertainties.

Thus, the gap between these regions may not be a reliable measure of the conservatism of the results. Note, however, that the design approach appears to provide more stability than is guaranteed a priori. Much of this conservatism may be attributable to the desire for a symmetric stability interval so close to an unstabilizable plant perturbation, i.e., $\sigma_1 = -1$. Nevertheless, the stability design objectives have been met in accordance with Theorem 6.2. Interestingly, the form of the actual stability region mimics the classical 6 dB downward/infinite dB upward gain margin of full-state-feedback LQR controllers ([1]). Thus, this approach appears to provide an alternative to gain-margin recovery techniques ([9]) which address this specialized form of plant uncertainty. Finally, Figure 2 compares guaranteed closed-loop performance to actual closed-loop performance over the guaranteed closed-loop robust stability region. Again the "actual" region was determined for constant $\hat{\sigma}_1(\cdot)$. Controller gains are given in Table 1. Finally, it is interesting to note that higher order robust controllers were obtained for this example in [46] using the μ -function approach.

Acknowledgment. I wish to thank Jill M. Straehla for preparing the manuscript versions of this paper, Scott W. Greeley for carrying out the numerical computations, Dr. Wassim M. Haddad for several helpful suggestions, and the reviewers for numerous helpful comments.

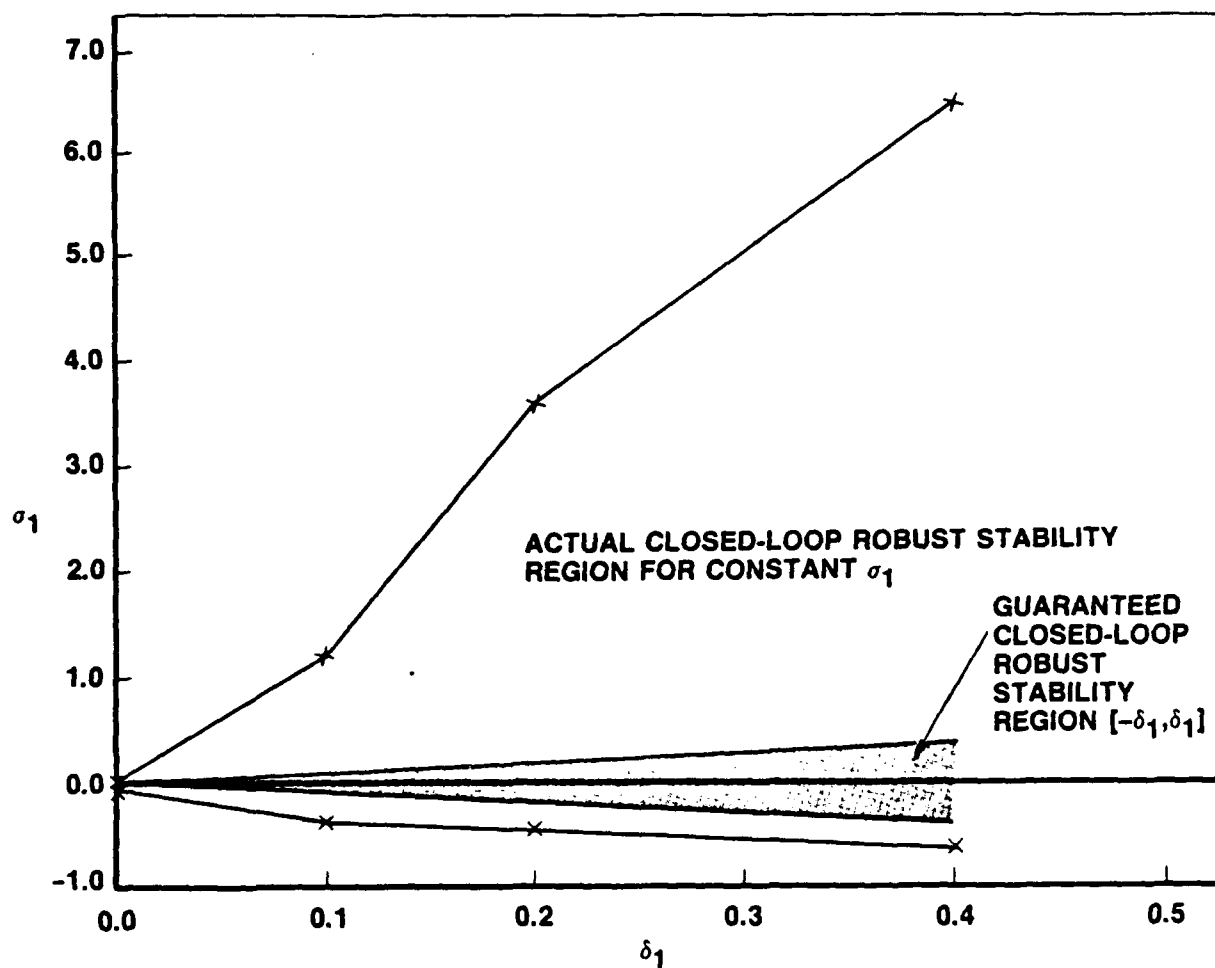


Figure 1

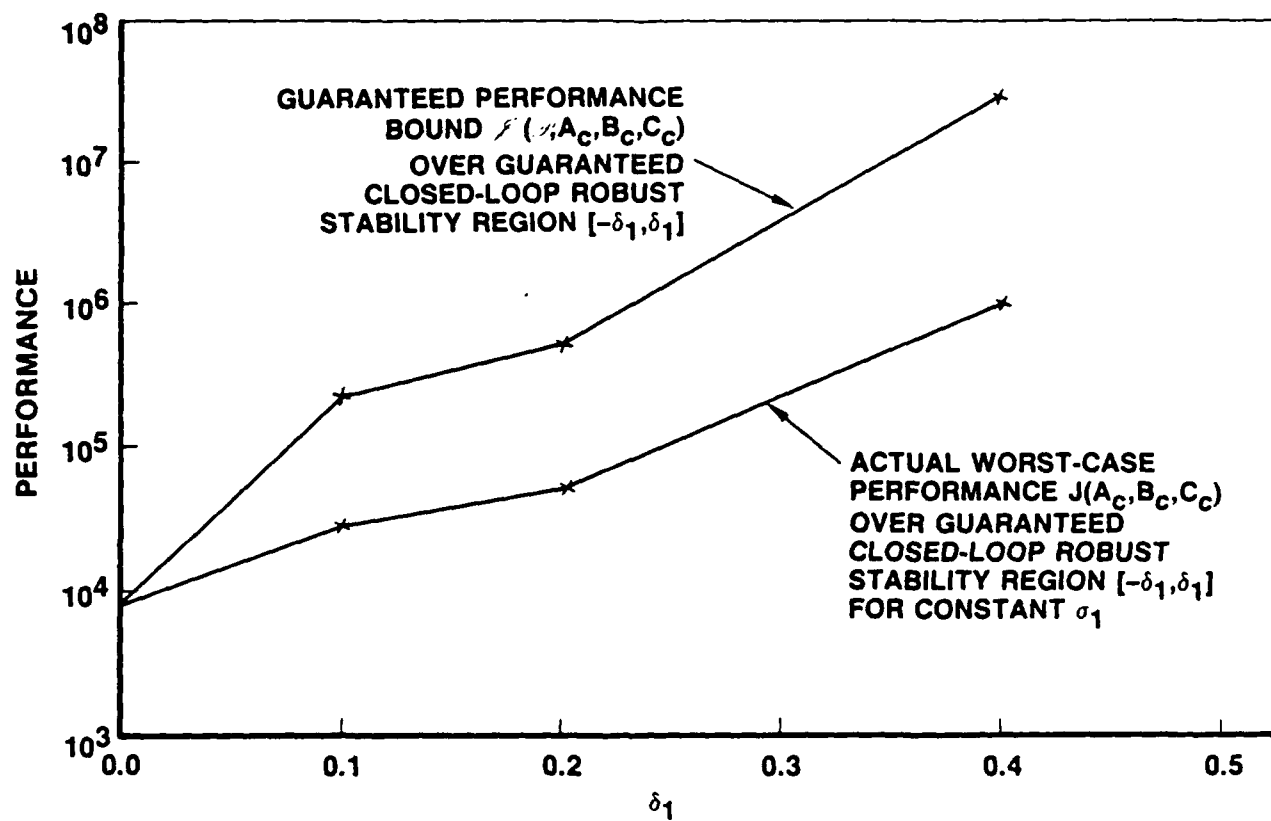


Figure 2

| (δ_1, α_1) | A_c | B_c | C_c |
|------------------------|--|--|---|
| $(.1, 1)$ | $\begin{bmatrix} -14.917 & 1.0 \\ -85.177 & 3.9657 \end{bmatrix}$ | $\begin{bmatrix} 15.917 \\ 79.959 \end{bmatrix}$ | $\begin{bmatrix} -5.2182 & -4.9657 \end{bmatrix}$ |
| $(.2, 2)$ | $\begin{bmatrix} -17.963 & 1.0 \\ -133.65 & -4.4614 \end{bmatrix}$ | $\begin{bmatrix} 18.963 \\ 127.05 \end{bmatrix}$ | $\begin{bmatrix} -6.6011 & -5.4614 \end{bmatrix}$ |
| $(.4, 4)$ | $\begin{bmatrix} -47.813 & 1.0 \\ -1087.3 & -6.5463 \end{bmatrix}$ | $\begin{bmatrix} 48.813 \\ 1073.5 \end{bmatrix}$ | $\begin{bmatrix} -13.766 & -7.5463 \end{bmatrix}$ |

Table 1

REFERENCES

1. M. G. Safonov and M. Athans, "Gain and Phase Margin for Multiloop LQG Regulators," *IEEE Trans. Autom. Contr.*, Vol. AC-22, pp. 173-179, 1977.
2. J. C. Doyle, "Guaranteed Margins for LQG Regulators," *IEEE Trans. Autom. Contr.*, Vol. AC-23, pp. 756-757, 1978.
3. J. C. Doyle and G. Stein, "Robustness with Observers," *IEEE Trans. Autom. Contr.*, Vol. AC-24, pp. 607-611, 1979.
4. J. C. Doyle and G. Stein, "Multivariable Feedback Design: Concepts for a Classical/Modern Synthesis," *IEEE Trans. Autom. Contr.*, Vol. AC-26, pp. 4-16, 1981.
5. G. Zames, "Feedback and Optimal Sensitivity: Model Reference Transformations, Multiplicative Seminorms, and Approximate Inverses," *IEEE Trans. Autom. Contr.*, Vol. AC-26, pp. 301-320, 1981.
6. J. C. Doyle, "Analysis of Feedback Systems with Structured Uncertainties," *IEE Proc.*, Vol. 129, pp. 242-250, 1982.
7. M. G. Safonov, "Stability Margins of Diagonally Perturbed Multivariable Feedback Systems," *IEE Proc.*, Vol. 129, pp. 251-256, 1982.
8. G. Zames and B. A. Francis, "Feedback, Minimax Sensitivity, and Optimal Robustness," *IEEE Trans. Autom. Contr.*, Vol. AC-28, pp. 585-601, 1983.
9. G. Stein and M. Athans, "The LQG/LTR Procedure for Multivariable Feedback Control Design," *IEEE Trans. Autom. Contr.*, Vol. AC-32, pp. 105-114, 1987.
10. B. A. Francis, *A Course in H_∞ Control Theory*, Springer-Verlag, New York, NY, 1987.
11. E. Soroka and U. Shaked, "On the Robustness of LQ Regulators," *IEEE Trans. Autom. Contr.*, Vol. AC-29, pp. 664-665, 1984.
12. U. Shaked and E. Soroka, "On the Stability Robustness of the Continuous-Time LQG Optimal Control," *IEEE Trans. Autom. Contr.*, Vol. AC-30, pp. 1039-1043, 1985.
13. M. Tahk and J. L. Speyer, "Modeling of Parameter Variations and Asymptotic LQG Synthesis," *Proc. 24th IEEE Conf. Dec. Contr.*, pp. 1459-1465, Athens, Greece, December 1986.
14. S. S. L. Chang and T. K. C. Peng, "Adaptive Guaranteed Cost Control of Systems with Uncertain Parameters," *IEEE Trans. Autom. Contr.*, Vol. AC-17, pp. 474-483, 1972.
15. R. V. Patel, M. Toda and B. Sridhar, "Robustness of Linear Quadratic State Feedback Designs in the Presence of System Uncertainty," *IEEE Trans. Autom. Contr.*, Vol. AC-22, pp. 945-949, 1977.
16. G. Leitmann, "Guaranteed Asymptotic Stability for a Class of Uncertain Linear Dynamical Systems," *J. Optim. Thy. Appl.*, Vol. 27, pp. 99-106, 1979.
17. A. Vinkler and L. J. Wood, "Multistep Guaranteed Cost Control of Linear Systems with Uncertain Parameters," *J. Guid. Contr.*, Vol. 2, pp. 449-456, 1979.
18. M. Eslami and D. L. Russell, "On Stability with Large Parameter Variations: Stemming from the Direct Method of Lyapunov," *IEEE Trans. Autom. Contr.*, Vol. AC-25, pp. 1231-1234, 1980.
19. M. Corless and G. Leitmann, "Continuous State Feedback Guaranteeing Uniform Ultimate Boundedness for Uncertain Dynamical Systems," *IEEE Trans. Autom. Contr.*, Vol. AC-26, pp. 1139-1144, 1981.
20. J. S. Thorp and B. R. Barmish, "On Guaranteed Stability of Uncertain Linear Systems via Linear Control," *J. Optim. Thy. Appl.*, Vol. 35, pp. 559-579, 1981.

21. B. R. Barmish, M. Corless and G. Leitmann, "A New Class of Stabilizing Controllers for Uncertain Dynamical Systems," *SIAM J. Contr. Optim.*, Vol. 21, pp. 246-255, 1983.
22. B. R. Barmish, I. R. Petersen and A. Feuer, "Linear Ultimate Boundedness Control of Uncertain Dynamic Systems," *Automatica*, Vol. 19, pp. 523-532, 1983.
23. B. R. Barmish, "Necessary and Sufficient Conditions for Quadratic Stabilizability of an Uncertain Linear System," *J. Optim. Thy. Appl.*, Vol. 46, pp. 399-408, 1985.
24. R. K. Yedavalli, S. S. Banda and D. B. Ridgely, "Time-Domain Stability Robustness Measures for Linear Regulators," *J. Guid. Contr. Dyn.*, Vol. 8, pp. 520-524, 1985.
25. A. R. Galimidi and B. R. Barmish, "The Constrained Lyapunov Problem and its Application to Robust Output Feedback Stabilization," *IEEE Trans. Autom. Contr.*, Vol. AC-31, pp. 410-419, 1986.
26. I. R. Petersen and C. V. Hollot, "A Riccati Equation Approach to the Stabilization of Uncertain Systems," *Automatica*, Vol. 22, pp. 397-411, 1986.
27. G. Leitmann, E. P. Ryan and A. Steinberg, "Feedback Control of Uncertain Systems: Robustness with Respect to Neglected Actuator and Sensor Dynamics," *Int. J. Contr.*, Vol. 43, pp. 1243-1256, 1986.
28. O. I. Kosmidou and P. Bertrand, "Robust-Controller Design for Systems with Large Parameter Variations," *Int. J. Contr.*, Vol. 45, pp. 927-938, 1987.
29. D. C. Hyland and D. S. Bernstein, "The Optimal Projection Equations for Fixed-Order Dynamic Compensation," *IEEE Trans. Autom. Contr.*, Vol. AC-29, pp. 1034-1037, 1984.
30. D. S. Bernstein and D. C. Hyland, "The Optimal Projection Equations for Finite-Dimensional Fixed-Order Dynamic Compensation of Infinite-Dimensional Systems," *SIAM J. Contr. Optim.*, Vol. 23, pp. 122-151, 1986.
31. D. S. Bernstein, "The Optimal Projection Equations for Static and Dynamic Output Feedback: The Singular Case," *IEEE Trans. Autom. Contr.*, Vol. AC-32, pp. 1139-1143, 1987.
32. D. S. Bernstein and W. M. Haddad, "LQG Control With An H_∞ Performance Bound," submitted to *IEEE Trans. Autom. Contr.*.
33. D. S. Bernstein, "Robust Static and Dynamic Output Feedback Stabilization: Deterministic and Stochastic Perspectives," *IEEE Trans. Autom. Contr.*, Vol. AC-32, pp. 1076-1084, 1987.
34. Y. A. Phillis, "Controller Design of Systems with Multiplicative Noise," *IEEE Trans. Autom. Contr.*, Vol. AC-30, pp. 1017-1019, 1985.
35. D. S. Bernstein and S. W. Greeley, "Robust Controller Synthesis Using the Maximum Entropy Design Equations," *IEEE Trans. Autom. Contr.*, Vol. AC-31, pp. 362-364, 1986.
36. D. S. Bernstein and D. C. Hyland, "The Optimal Projection Equations for Reduced-Order Modelling, Estimation and Control of Linear Systems with Multiplicative White Noise," *J. Optim. Thy. Appl.*, Vol. 58, 1988, to appear.
37. S. Richter, "A Homotopy Algorithm for Solving the Optimal Projection Equations for Fixed-Order Dynamic Compensation: Existence, Convergence and Global Optimality," *Proc. Amer. Contr. Conf.*, pp. 1527-1531, Minneapolis, MN, June 1987.
38. S. W. Greeley and D. C. Hyland, "Reduced-Order Compensation: LQG Reduction Versus Optimal Projection Using a Homotopic Continuation Method," *Proc. IEEE Conf. Dec. Contr.*, pp. 742-747, Los Angeles, CA, December 1987.
39. A. Gruzen, "Robust Reduced Order Control of Flexible Structures," C. S. Draper Laboratory Report #CSDL-T-900, April 1986.

40. A. Gruzen and W. E. Vander Velde, "Robust Reduced-Order Control of Flexible Structures Using the Optimal Projection/Maximum Entropy Design Methodology," *AIAA Guid. Nav. Contr. Conf.*, Williamsburg, VA, August 1978.
41. J. W. Brewer, "Kronecker Products and Matrix Calculus in System Theory," *IEEE Trans. Circ. Sys.*, Vol. CAS-25, pp. 772-781, 1978.
42. J. K. Hale, *Ordinary Differential Equations*, Wiley, New York, 1969.
43. J. Warga, *Optimal Control of Differential and Functional Equations*, Academic Press, New York, 1972.
44. L. Arnold, *Stochastic Differential Equations: Theory and Applications*, Wiley, New York, 1974.
45. C. R. Rao and S. K. Mitra, *Generalized Inverse of Matrices and Its Applications*, John Wiley and Sons, New York, 1971.
46. A. Packard and J. C. Doyle, "Robust Control with an H_2 Performance Objective," *Proc. Amer. Contr. Conf.*, pp. 2141-2145, Minneapolis, MN, June 1987.

APPENDIX J: H_∞ Theory

- 114. W. M. Haddad and D. S. Bernstein, "Combined L_2/H_∞ Model Reduction," Int. J. Contr., to appear.
- 116. D. S. Bernstein and W. M. Haddad, "Steady-State Kalman Filtering with an H_∞ Error Bound," Sys. Contr. Lett., Vol. 12, 1989, to appear.
- 117. D. S. Bernstein and W. M. Haddad, "LQG Control With An H_∞ Performance Bound: A Riccati Equation Approach," IEEE Trans. Autom. Contr., Vol. 34, 1989, to appear.
- 105. A. N. Madiwale, W. M. Haddad, and D. S. Bernstein, "Robust H_∞ Controllers for Systems with Parameter Uncertainty," Proc. IEEE Conf. Dec. Contr., Austin, TX, December 1988.

February 1988
Revised April 1988

Combined L_2/H_∞ Model Reduction

by

Wassim M. Haddad
Dept. of Mechanical Engineering
Florida Institute of Technology
Melbourne, FL 32901

Dennis S. Bernstein
Harris Corporation
Government Aerospace Systems Division
MS 22/4848
Melbourne, FL 32902

Abstract

A model-reduction problem is considered which involves both L_2 (quadratic) and H_∞ (worst-case frequency-domain) aspects. Specifically, the goal of the problem is to minimize an L_2 model-reduction criterion subject to a prespecified H_∞ constraint on the model-reduction error. The principal result is a sufficient condition for characterizing reduced-order models with bounded L_2 and H_∞ approximation error. The sufficient condition involves a system of modified Riccati equations coupled by an oblique projection, i.e., idempotent matrix. When the H_∞ constraint is absent, the sufficient condition specializes to the L_2 model-reduction result given in Hyland and Bernstein, 1985.

Supported in part by the Air Force Office of Scientific Research under contract F49620-86-C-0002

1. Introduction

One of the most fundamental problems in dynamic systems theory is to approximate a high-order, complex system with a low-order, relatively simpler model. The resulting reduced-order model can then be used to facilitate the analysis of complex systems as well as the design and implementation of feedback controllers and electronic filters. The model-reduction problem thus reflects the fundamental engineering desire for simplicity of implementation and parsimony of hardware.

In view of the practical motivations for the model-reduction problem, it is not surprising that significant effort has been devoted to this problem in recent years. Indeed, there now exists a well-developed theoretical foundation for model reduction under a variety of approximation criteria. Expanding on the original work of Adamjan, Arov and Krein, 1971, progress was achieved in Kung and Lin, 1981, Lin and Kung, 1982, Glover, 1984, Latham and Anderson, 1985, Hung and Glover, 1986, Anderson, 1986, Ball and Ran, 1987, and Parker and Anderson, 1987, for the Hankel-norm approximation criterion. Many of the cited works also present bounds for the closely related H_∞ approximation error, although the optimal H_∞ model-reduction problem remains open. Alternatively, early progress on the model-reduction problem with a quadratic (L_2) criterion was achieved in Wilson, 1970, and further explored in Hyland and Bernstein, 1985.

Although the Hankel norm, H_∞ , and L_2 model-reduction criteria represent distinct approximation objectives, there exist significant connections. For example, it was shown in Wilson, 1985, that for systems which are either single input or single output, the input and output space topologies can be redefined so that the induced norm of the Hankel operator coincides with the L_2 system norm. In addition, the optimization technique utilized in Wilson, 1970, was reapplied to the Hilbert-Schmidt Hankel operator topology in Wilson, 1988. In recent work, Wilson, 1988a, has shown that for single-input or single-output systems the quadratic model-reduction criterion is actually an induced norm of the convolution operator itself.

In the present paper we attempt a further unification of the L_2 and H_∞ model-reduction objectives. Specifically, we consider an L_2 model-reduction problem with a constraint on the H_∞ approximation error. The underlying idea involves the suitable application of a frequency-domain inequality due to Willems, 1971, which has recently been applied to H_∞ control-design problems in Petersen, 1987, Khargonekar, Petersen and Zhou, 1987, and Bernstein and Haddad, 1988. The principle result of the present paper is a sufficient condition which characterizes reduced-order

models satisfying an optimized L_2 bound as well as a prespecified H_∞ bound. The sufficient condition is a direct generalization of the optimal projection approach developed in Hyland and Bernstein, 1985, for the unconstrained L_2 problem. While the L_2 -optimal reduced-order model was characterized in Hyland and Bernstein, 1985, by means of a coupled system of two modified Lyapunov equations, the H_∞ -constrained solution in the present paper involves a coupled system consisting of four modified Riccati equations. As in Hyland and Bernstein, 1985, the coupling is due to the presence of an oblique projection (idempotent matrix) which determines the constrained reduced-order model. When the H_∞ constraint is sufficiently relaxed, we show that the conditions given herein specialize directly to those given in Hyland and Bernstein, 1985. Although our result gives sufficient conditions for H_∞ approximation, we also state hypotheses under which these conditions are also necessary.

Although numerical algorithms were developed in Hyland and Bernstein, 1985, for the "pure" L_2 problem, computational methods for the H_∞ -constrained problem are beyond the scope of the present paper. In view of the additional complexity engendered by the H_∞ constraint, more sophisticated algorithms appear necessary. Hence computational methods will focus on the homotopic continuation algorithm developed in Richter, 1987, for reduced-order dynamic compensation.

Notation and Definitions

| | |
|---|--|
| $\mathbb{R}, \mathbb{R}^{r \times s}, \mathbb{R}^r, \mathbb{E}$ | real numbers, $r \times s$ real matrices, $\mathbb{R}^{r \times 1}$, expected value |
| $I_r, ()^T, 0_{r \times s}, 0_r$ | $r \times r$ identity matrix, transpose, $r \times s$ zero matrix, $0_{r \times r}$ |
| $()^*$ | complex conjugate transpose |
| tr | trace |
| $\sigma_{\max}(Z)$ | largest singular value of matrix Z |
| $\lambda_{\max}(Z)$ | largest eigenvalue of matrix Z with real spectrum |
| $\ Z\ _F$ | $[\text{tr } ZZ^*]^{\frac{1}{2}}$ (Frobenius matrix norm) |
| $\ h(t)\ _2$ | $[\int_0^\infty \ h(t)\ _F^2 dt]^{\frac{1}{2}}$ |
| $\ H(s)\ _2$ | $[\frac{1}{2\pi} \int_{-\infty}^\infty \ H(j\omega)\ _F^2 d\omega]^{\frac{1}{2}}$ |
| $\ H(s)\ _\infty$ | $\sup_{\omega \in \mathbb{R}} \sigma_{\max}[H(j\omega)]$ |
| $\mathbb{S}^r, \mathbb{N}^r, \mathbb{P}^r$ | $r \times r$ symmetric, nonnegative-definite, positive-definite matrices |
| $Z_1 \leq Z_2, Z_1 < Z_2$ | $Z_2 - Z_1 \in \mathbb{N}^r, Z_2 - Z_1 \in \mathbb{P}^r, Z_1, Z_2 \in \mathbb{S}^r$ |
| $n, m, \ell, n_m, q, p; \tilde{n}$ | positive integers; $n + n_m$ |
| $x, y, y_m, x_m, \tilde{y}, \tilde{x}$ | $n, \ell, \ell, n_m, \ell, \tilde{n}$ -dimensional vectors |

| | |
|-----------------------------------|--|
| \tilde{y}, \tilde{x} | $y - y_m, \begin{bmatrix} x \\ x_m \end{bmatrix}$ |
| A, B, C | $n \times n, n \times m, \ell \times n$ matrices |
| D, E | $m \times p, q \times \ell$ matrices |
| A_m, B_m, C_m | $n_m \times n_m, n_m \times m, \ell \times n_m$ matrices |
| $\tilde{A}, \tilde{B}, \tilde{C}$ | $\begin{bmatrix} A & 0 \\ 0 & A_m \end{bmatrix}, \begin{bmatrix} B \\ B_m \end{bmatrix}, [C \quad -C_m]$ |
| \tilde{D}, \tilde{E} | $\tilde{B}D = \begin{bmatrix} BD \\ B_m D \end{bmatrix}, E\tilde{C} = [EC \quad -EC_m]$ |
| R | $E^T E$, model-reduction error-weighting matrix in \mathbb{P}^ℓ |
| $w(\cdot)$ | p -dimensional standard white noise process |
| V | intensity of $Dw(\cdot)$, $V = DD^T \in \mathbb{P}^m$ |
| \tilde{R}, \tilde{V} | $\begin{bmatrix} C^T R C & -C^T R C_m \\ -C_m^T R C & C_m^T R C_m \end{bmatrix}, \begin{bmatrix} B V B^T & B V B_m^T \\ B_m V B^T & B_m V B_m^T \end{bmatrix}$ |
| γ | positive constant |

2. Statement of the Problem

In this section we introduce the model-reduction problem with constrained H_∞ norm of the model-reduction error. Specifically, we constrain the transfer function of the reduced-order model to lie within a specified H_∞ radius of the original system. In this paper we assume that the full-order model is asymptotically stable, i.e., the matrix A is asymptotically stable.

H_∞ -Constrained L_2 Model-Reduction Problem. Given the n th-order controllable and observable model

$$\dot{x}(t) = Ax(t) + BDw(t), \quad (2.1)$$

$$y(t) = Cx(t), \quad (2.2)$$

where $t \in [0, \infty)$, determine an n_m th-order model

$$\dot{x}_m(t) = A_m x_m(t) + B_m Dw(t), \quad (2.3)$$

$$y_m(t) = C_m x_m(t), \quad (2.4)$$

which satisfies the following criteria:

- (i) A_m is asymptotically stable;

- (ii) the transfer function of the reduced-order model is within a radius- γ H_∞ neighborhood of the full-order model, i.e.,

$$\|H(s) - H_m(s)\|_\infty \leq \gamma, \quad (2.5)$$

where

$$H(s) \triangleq EC(sI_n - A)^{-1}BD, \quad H_m(s) \triangleq EC_m(sI_{n_m} - A_m)^{-1}B_mD, \quad (2.6)$$

and $\gamma > 0$ is a given constant; and

- (iii) the L_2 model-reduction criterion

$$J(A_m, B_m, C_m) \triangleq \lim_{t \rightarrow \infty} \mathbb{E} \left\{ [y(t) - y_m(t)]^T R [y(t) - y_m(t)] \right\} \quad (2.7)$$

is minimized.

Note that the full- and reduced-order systems (2.1)–(2.4) can be written as a single augmented system

$$\dot{\tilde{x}}(t) = \tilde{A}\tilde{x}(t) + \tilde{D}w(t), \quad t \in [0, \infty), \quad (2.8)$$

so that the $q \times p$ transfer function from $w(t)$ to $E\tilde{y}(t) = \tilde{E}\tilde{x}(t)$ is

$$\tilde{H}(s) = \tilde{E}(sI_{\tilde{n}} - \tilde{A})^{-1}\tilde{D} \quad (2.9)$$

and (2.7) can be written as

$$J(A_m, B_m, C_m) = \lim_{t \rightarrow \infty} \mathbb{E} \left\{ [E\tilde{y}(t)]^T [E\tilde{y}(t)] \right\} = \lim_{t \rightarrow \infty} \mathbb{E} [\tilde{x}^T(t) \tilde{R} \tilde{x}(t)]. \quad (2.10)$$

Before continuing it is useful to note that if A_m is asymptotically stable then the L_2 model-reduction criterion (2.7) is given by

$$J(A_m, B_m, C_m) = \text{tr } \tilde{Q} \tilde{R}, \quad (2.11)$$

where the steady-state covariance

$$\tilde{Q} \triangleq \lim_{t \rightarrow \infty} \mathbb{E} [\tilde{x}(t) \tilde{x}^T(t)] \quad (2.12)$$

satisfies the augmented Lyapunov equation

$$0 = \tilde{A}\tilde{Q} + \tilde{Q}\tilde{A}^T + \tilde{V}. \quad (2.13)$$

Using (2.11) and (2.13) it can be shown that the L_2 criterion (2.7) is an approximation measure involving the full- and reduced-order impulse responses with respect to an L_2 norm.

Proposition 2.1. The L_2 model-reduction criterion (2.11) can be written as

$$J(A_m, B_m, C_m) = \int_0^\infty \|ECe^{A_m t}BD - EC_me^{A_m t}B_mD\|_F^2 dt, \quad (2.14a)$$

or, equivalently,

$$J(A_m, B_m, C_m) = \frac{1}{2\pi} \int_{-\infty}^\infty \|H(j\omega) - H_m(j\omega)\|_F^2 d\omega. \quad (2.14b)$$

Proof. It need only be noted that (2.11) is equivalent to

$$\begin{aligned} \text{tr} \int_0^\infty e^{\tilde{A}t} \tilde{V} e^{\tilde{A}^T t} dt \tilde{R} &= \text{tr} \int_0^\infty \tilde{E} e^{\tilde{A}t} \tilde{D} \tilde{D}^T e^{\tilde{A}^T t} \tilde{E}^T dt \\ &= \text{tr} \int_0^\infty (\tilde{E} e^{\tilde{A}t} \tilde{D})(\tilde{E} e^{\tilde{A}t} \tilde{D})^T dt \\ &= \int_0^\infty \|\tilde{E} e^{\tilde{A}t} \tilde{D}\|_F^2 dt \end{aligned}$$

which is equivalent to (2.14a). Finally, (2.14b) follows from Plancherel's Theorem. \square

The key step in enforcing (2.5) is to replace the algebraic Lyapunov equation (2.13) by an algebraic Riccati equation. Justification for this technique is provided by the following result.

Lemma 2.1. Let (A_m, B_m, C_m) be given and assume there exists $Q \in \mathbb{R}^{\tilde{n} \times \tilde{n}}$ satisfying

$$Q \in \mathbb{N}^{\tilde{n}} \quad (2.15)$$

and

$$0 = \tilde{A}Q + Q\tilde{A}^T + \gamma^{-2}Q\tilde{R}Q + \tilde{V}. \quad (2.16)$$

Then

$$(\tilde{A}, [\gamma^{-2}Q\tilde{R}Q + \tilde{V}]^{\frac{1}{2}}) \text{ is stabilizable} \quad (2.17)$$

if and only if

$$A_m \text{ is asymptotically stable.} \quad (2.18)$$

Furthermore, in this case,

$$\|H(s) - H_m(s)\|_\infty \leq \gamma, \quad (2.19)$$

$$\tilde{Q} \leq Q, \quad (2.20)$$

and

$$J(A_m, B_m, C_m) \leq J(A_m, B_m, C_m, Q), \quad (2.21)$$

where

$$J(A_m, B_m, C_m, Q) \triangleq \text{tr } Q\tilde{R}. \quad (2.22)$$

Proof. Using the assumed existence of a nonnegative-definite solution to (2.16) and the stabilizability condition (2.17), it follows from the dual of Lemma 12.2 of Wonham, 1979, that \tilde{A} is asymptotically stable. Since \tilde{A} is block diagonal, A_m is also asymptotically stable. Conversely, since A is assumed to be asymptotically stable, (2.18) implies (2.17). The proof of (2.19) follows from a standard manipulation of (2.16); for details see Lemma 1 of Willems, 1971. To prove (2.20) subtract (2.13) from (2.16) to obtain

$$0 = \tilde{A}(Q - \tilde{Q}) + (Q - \tilde{Q})\tilde{A}^T + \gamma^{-2}Q\tilde{R}Q, \quad (2.23)$$

which, since \tilde{A} is asymptotically stable, is equivalent to

$$Q - \tilde{Q} = \int_0^\infty e^{\tilde{A}t} [\gamma^{-2}Q\tilde{R}Q] e^{\tilde{A}^T t} dt \geq 0. \quad (2.24)$$

Finally, (2.21) follows immediately from (2.20). \square

Lemma 2.1 shows that the H_∞ constraint is automatically enforced when a nonnegative-definite solution to (2.16) is known to exist. Furthermore, the solution Q provides an upper bound for the actual state covariance \tilde{Q} along with a bound on the L_2 model-reduction criterion. Next, we present a partial converse of Lemma 2.1 which guarantees the existence of a nonnegative-definite solution to (2.16) when (2.19) is satisfied.

Lemma 2.2. Let (A_m, B_m, C_m) be given, suppose \tilde{A} is asymptotically stable, and assume the H_∞ approximation constraint (2.19) is satisfied. Then there exists a unique nonnegative-definite solution Q satisfying (2.16). Furthermore, $(\tilde{A} + \gamma^{-2}Q\tilde{R}, \tilde{V}^{\frac{1}{2}})$ is stabilizable if and only if $\tilde{A} + \gamma^{-2}Q\tilde{R}$ is asymptotically stable.

Proof. The result is an immediate consequence of Theorems 3 and 2, pp. 150 and 167 of Brockett, 1970, and the dual of Lemma 12.2 of Wonham, 1979. \square

Finally, we show that the quadratic term $\gamma^{-2}Q\tilde{R}Q$ in (2.16) also constrains the Hankel norm of the approximation error $E\tilde{y}$ when Q is positive definite. To show this let $\tilde{P} \in \mathbb{N}^n$ be the

observability Gramian for the augmented system $(\tilde{A}, \tilde{D}; \tilde{E})$ which satisfies

$$0 = \tilde{A}^T \tilde{P} + \tilde{P} \tilde{A} + \tilde{R}. \quad (2.25)$$

Furthermore, note that \tilde{Q} satisfying (2.13) is the dual controllability Gramian.

Proposition 2.2. Let (A_m, B_m, C_m) be given and assume there exists $Q \in \mathbb{P}^{\tilde{n}}$ satisfying (2.16) and (2.17) or, equivalently, (2.18). Then

$$\lambda_{\max}^{\frac{1}{2}}(\tilde{P}\tilde{Q}) \leq \gamma. \quad (2.26)$$

Proof. Since Q is invertible, (2.16) implies

$$0 = \gamma^2 \tilde{A}^T Q^{-1} + \gamma^2 Q^{-1} \tilde{A} + \gamma^2 Q^{-1} \tilde{V} Q^{-1} + \tilde{R}. \quad (2.27)$$

Next, subtract (2.25) from (2.27) to obtain

$$0 = \tilde{A}^T (\gamma^2 Q^{-1} - \tilde{P}) + (\gamma^2 Q^{-1} - \tilde{P}) \tilde{A} + \gamma^2 Q^{-1} \tilde{V} Q^{-1}, \quad (2.28)$$

which, since \tilde{A} is asymptotically stable, is equivalent to

$$\gamma^2 Q^{-1} - \tilde{P} = \int_0^\infty e^{\tilde{A}^T t} [\gamma^2 Q^{-1} \tilde{V} Q^{-1}] e^{\tilde{A} t} dt \geq 0. \quad (2.29)$$

Thus, (2.29) implies $\tilde{P} \leq \gamma^2 Q^{-1}$ or, equivalently, $Q^{\frac{1}{2}} \tilde{P} Q^{\frac{1}{2}} \leq \gamma^2 I_{\tilde{n}}$. Hence, $\lambda_{\max}^{\frac{1}{2}}(\tilde{P}Q) \leq \gamma$. Finally, (2.26) follows immediately from (2.20). \square

3. The Auxiliary Minimization Problem and Necessary Conditions for Optimality

As discussed in the previous section, the replacement of (2.13) by (2.16) enforces the H_∞ approximation constraint between the full- and reduced-order systems and results in an upper bound for the L_2 model-reduction criterion. That is, if (2.16) is solvable then the reduced-order model (A_m, B_m, C_m) satisfies the H_∞ approximation constraint (2.5) while the *actual* L_2 model-reduction criterion is guaranteed to be no worse than the bound given by $J(A_m, B_m, C_m, Q)$. Hence, $J(A_m, B_m, C_m, Q)$ can be interpreted as an *auxiliary cost* which leads to the following mathematical programming problem.

Auxiliary Minimization Problem. Determine (A_m, B_m, C_m, Q) which minimizes $J(A_m, B_m, C_m, Q)$ subject to (2.15) and (2.16).

It follows from Lemma 2.1 that the satisfaction of (2.15)–(2.17) leads to 1) A_m stable; 2) a bound on the H_∞ distance between the full-order and reduced-order systems; and 3) an upper bound for the L_2 model-reduction criterion. Hence it remains to determine (A_m, B_m, C_m) which minimizes $J(A_m, B_m, C_m, Q)$ and thus provides an optimized bound for the *actual* L_2 criterion $J(A_m, B_m, C_m)$. Rigorous derivation of the necessary conditions for the Auxiliary Minimization Problem requires additional technical assumptions. Specifically, we restrict (A_m, B_m, C_m, Q) to the open set

$$S \triangleq \{(A_m, B_m, C_m, Q) : Q \in \mathbb{P}^n, \tilde{A} + \gamma^{-2} Q \tilde{R} \text{ is asymptotically stable,} \\ \text{and } (A_m, B_m, C_m) \text{ is controllable and observable}\}. \quad (3.1)$$

Remark 3.1. The set S constitutes sufficient conditions under which the Lagrange multiplier technique is applicable to the Auxiliary Minimization Problem. Specifically, the requirement that Q be positive definite replaces (2.15) by an open set constraint, the stability of $\tilde{A} + \gamma^{-2} Q \tilde{R}$ serves as a normality condition, and (A_m, B_m, C_m) minimal is a nondegeneracy condition.

The following Lemma is needed for the statement of the main result.

Lemma 3.1. Let $\hat{Q}, \hat{P} \in \mathbb{P}^n$ and suppose $\text{rank } \hat{Q}\hat{P} = n_m$. Then there exist $n_m \times n$ G, Γ and $n_m \times n_m$ invertible M , unique except for a change of basis in \mathbb{R}^{n_m} , such that

$$\hat{Q}\hat{P} = G^T M \Gamma, \quad (3.2)$$

$$\Gamma G^T = I_{n_m}. \quad (3.3)$$

Furthermore, the $n \times n$ matrices

$$r \triangleq G^T \Gamma, \quad (3.4)$$

$$r_\perp \triangleq I_n - r \quad (3.5)$$

are idempotent and have rank n_m and $n - n_m$, respectively. If, in addition,

$$\text{rank } \hat{Q} = \text{rank } \hat{P} = n_m, \quad (3.6)$$

then

$$\hat{Q} = r \hat{Q}, \quad \hat{P} = \hat{P} r. \quad (3.7), (3.8)$$

Finally, if $P \in \mathbb{P}^n$ then the inverse

$$S \triangleq (I_n + \gamma^{-2} \hat{Q} P)^{-1} \quad (3.9)$$

exists.

Proof. Conditions (3.2)–(3.8) are a direct consequence of Theorem 6.2.5 of Rao and Mitra, 1971. To prove that the inverse in (3.9) exists, note that since the eigenvalues of $\hat{Q}P$ coincide with the eigenvalues of the nonnegative-definite matrix $P^{\frac{1}{2}}\hat{Q}P^{\frac{1}{2}}$, it follows that $\hat{Q}P$ has nonnegative eigenvalues. Thus, the eigenvalues of $I_n + \gamma^{-2}\hat{Q}P$ are all greater than one so that the above inverse exists. \square

Finally, for convenience define

$$\Sigma \triangleq BV B^T, \quad \bar{\Sigma} \triangleq C^T R C.$$

Theorem 3.1. If $(A_m, B_m, C_m, Q) \in S$ solves the Auxiliary Minimization Problem then there exist $Q, P, \hat{Q}, \hat{P} \in \mathbb{N}^n$ such that

$$A_m = \Gamma(A - \gamma^{-4}Q\bar{\Sigma}QPS)G^T, \quad (3.10)$$

$$B_m = \Gamma B, \quad (3.11)$$

$$C_m = C(I_n + \gamma^{-2}QPS)G^T, \quad (3.12)$$

$$Q = \begin{bmatrix} Q + \hat{Q} & \hat{Q}\Gamma^T \\ \Gamma\hat{Q} & \Gamma\hat{Q}\Gamma^T \end{bmatrix}, \quad (3.13)$$

and such that Q, P, \hat{Q}, \hat{P} satisfy

$$0 = AQ + QA^T + \gamma^{-2}Q\bar{\Sigma}Q + r_{\perp}\Sigma r_{\perp}^T, \quad (3.14)$$

$$0 = A^T P + PA - \gamma^{-4}S^T P Q \bar{\Sigma} Q P S + r_{\perp}^T (I_n + \gamma^{-2}QPS)^T \bar{\Sigma} (I_n + \gamma^{-2}QPS) r_{\perp}, \quad (3.15)$$

$$0 = (A - \gamma^{-4}Q\bar{\Sigma}QPS)\hat{Q} + \hat{Q}(A - \gamma^{-4}Q\bar{\Sigma}QPS)^T + \gamma^{-4}\hat{Q}S^T P Q \bar{\Sigma} Q P S \hat{Q} + \Sigma - r_{\perp}\Sigma r_{\perp}^T, \quad (3.16)$$

$$0 = (A + \gamma^{-2}Q\bar{\Sigma})^T \hat{P} + \hat{P}(A + \gamma^{-2}Q\bar{\Sigma}) + (I_n + \gamma^{-2}QPS)^T \bar{\Sigma} (I_n + \gamma^{-2}QPS) - r_{\perp}^T (I_n + \gamma^{-2}QPS)^T \bar{\Sigma} (I_n + \gamma^{-2}QPS) r_{\perp}, \quad (3.17)$$

$$\text{rank } \hat{Q} = \text{rank } \hat{P} = \text{rank } \hat{Q}\hat{P} = n_m. \quad (3.18)$$

Furthermore, the auxiliary cost is given by

$$J(A_m, B_m, C_m, Q) = \text{tr } \bar{\Sigma}(Q + \gamma^{-4}QPS\hat{Q}S^T P Q). \quad (3.19)$$

Conversely, if there exist $Q, P, \hat{Q}, \hat{P} \in \mathbb{N}^n$ satisfying (3.14)–(3.18), then (A_m, B_m, C_m, Q) given by (3.10)–(3.13) satisfy (2.15) and (2.16) with the auxiliary cost (2.22) given by (3.19).

Proof. See Appendix A. \square

Remark 3.1. Theorem 3.1 presents necessary conditions for the Auxiliary Minimization Problem which explicitly synthesize extremal reduced-order models (A_m, B_m, C_m) . As a check of these conditions, consider the extreme case $n_m = n$. Then $G = \Gamma^{-1}$ and thus, without loss of generality, $G = \Gamma = r = I_n$ and $r_\perp = 0$. Furthermore, (3.14) implies that $Q = 0$ and (3.15) implies that $P = 0$. Hence the H_∞ -constrained full-order model is given (as expected) by (A, B, C) regardless of γ . Furthermore, note that \mathcal{Q} given by (3.13) becomes

$$\mathcal{Q} = \begin{bmatrix} \hat{Q} & \hat{Q} \\ \hat{Q} & \hat{Q} \end{bmatrix} \quad (3.20)$$

so that the quadratic term $\gamma^{-2} \mathcal{Q} \tilde{R} \mathcal{Q}$ in (2.16) vanishes. Thus (2.16) reduces to (2.13) so that \mathcal{Q} coincides with the controllability Gramian \tilde{Q} . If, alternatively, the reduced-order constraint is retained but the transfer function approximation constraint (2.5) is sufficiently relaxed, i.e., $\gamma \rightarrow \infty$, then $S = I_n$ so that the reduced-order model (3.10)–(3.12) is given by $(A_m, B_m, C_m) = (\Gamma A \Gamma^T, \Gamma B, C \Gamma^T)$. In this case (3.14) and (3.15) are superfluous and (3.16) and (3.17) reduce to the optimal projection equations obtained by Hyland and Bernstein, 1985, for the unconstrained L_2 problem.

4. Sufficient Conditions for Combined L_2/H_∞ Approximation

In this section we combine Lemma 2.1 with the converse of Theorem 3.1 to obtain our main result guaranteeing constrained H_∞ approximation along with an optimized L_2 model-reduction bound.

Theorem 4.1. Suppose there exist $Q, P, \hat{Q}, \hat{P} \in \mathbb{R}^n$ satisfying (3.14)–(3.18) and let $(A_m, B_m, C_m, \mathcal{Q})$ be given by (3.10)–(3.13). Then $(\tilde{A}, [\gamma^{-2} \mathcal{Q} \tilde{R} \mathcal{Q} + \tilde{V}]^{\frac{1}{2}})$ is stabilizable if and only if A_m is asymptotically stable. In this case, the reduced-order transfer function $H_m(s)$ satisfies the H_∞ approximation constraint

$$\|H(s) - H_m(s)\|_\infty \leq \gamma \quad (4.1)$$

and the L_2 approximation bound

$$\|H(s) - H_m(s)\|_2 \leq [\text{tr } \tilde{E}(Q + \gamma^{-4} Q P S \hat{Q} S^T P Q)]^{\frac{1}{2}}. \quad (4.2)$$

Proof. The converse portion of Theorem 3.1 implies that \mathcal{Q} given by (3.13) satisfies (2.15) and (2.16) with auxiliary cost given by (3.19). It now follows from Lemma 2.1 that the stabilizability

condition (2.17) is equivalent to the asymptotic stability of A_m , the H_∞ approximation condition (2.19) holds, and the L_2 model-reduction criterion satisfies the bound (2.21) which is equivalent to (4.2). \square

In applying Theorem 4.1 the principal issue concerns conditions on the problem data under which the coupled Riccati equations (3.14)–(3.17) possess nonnegative-definite solutions. Clearly, for γ sufficiently large, (3.14)–(3.17) approximate the "pure" L_2 solution obtained in Hyland and Bernstein, 1985. In practice, we would numerically solve (3.14)–(3.17) for successively smaller values of γ until solutions are no longer obtainable. The important case of interest, however, involves small γ so that accurate H_∞ approximation is enforced. Thus, if (4.1) can be satisfied for a given $\gamma > 0$ by a class of reduced-order models, it is of interest to know whether one such reduced-order model can be obtained by solving (3.14)–(3.17). Lemma 2.2 guarantees that (2.16) possesses a solution for any model satisfying (4.1). Thus our sufficient conditions will also be necessary so long as the Auxiliary Minimization Problem possesses at least one extremal over S . When this is the case we have the following immediate result.

Proposition 4.1. Let γ^* denote the infimum of $\|H(s) - H_m(s)\|_\infty$ over all asymptotically stable reduced-order models and suppose that the Auxiliary Minimization Problem has a solution for all $\gamma > \gamma^*$. Then for all $\gamma > \gamma^*$ there exist $Q, P, \hat{Q}, \hat{P} \in \mathbb{N}^n$ satisfying (3.14)–(3.17).

Remark 4.1. As in Hyland and Bernstein, 1985, it can be expected that (3.14)–(3.17) possess multiple solutions. Theorem 4.1 guarantees, however, that the bounds (4.1) and (4.2) are enforced for all such extremals obtained by solving (3.14)–(3.17).

Appendix A: Proof of Theorem 3.1

To optimize (2.22) over the open set S subject to the constraint (2.16), form the Lagrangian

$$\mathcal{L}(A_m, B_m, C_m, Q, P, \lambda) \triangleq \text{tr}\{\lambda Q \bar{R} + [\bar{A}Q + Q\bar{A}^T + \gamma^{-2}Q\bar{R}Q + \bar{V}]P\}, \quad (\text{A.1})$$

where the Lagrange multipliers $\lambda \geq 0$ and $P \in \mathbb{R}^{n \times n}$ are not both zero. We thus obtain

$$\frac{\partial \mathcal{L}}{\partial Q} = (\bar{A} + \gamma^{-2}Q\bar{R})^T P + P(\bar{A} + \gamma^{-2}Q\bar{R}) + \lambda \bar{R}. \quad (\text{A.2})$$

Setting $\frac{\partial \mathcal{L}}{\partial Q} = 0$ yields

$$0 = (\bar{A} + \gamma^{-2}Q\bar{R})^T P + P(\bar{A} + \gamma^{-2}Q\bar{R}) + \lambda \bar{R}. \quad (\text{A.3})$$

Since $\bar{A} + \gamma^{-2}Q\bar{R}$ is assumed to be stable, $\lambda = 0$ implies $P = 0$. Hence, it can be assumed without loss of generality that $\lambda = 1$. Furthermore, P is nonnegative definite.

Now partition $\bar{n} \times \bar{n}$ Q, P , into $n \times n, n \times n_m$, and $n_m \times n_m$ subblocks as

$$Q = \begin{bmatrix} Q_1 & Q_{12} \\ Q_{12}^T & Q_2 \end{bmatrix}, \quad P = \begin{bmatrix} P_1 & P_{12} \\ P_{12}^T & P_2 \end{bmatrix},$$

and for notational convenience define

$$PQ = \begin{bmatrix} Z_1 & Z_{12} \\ Z_{21} & Z_2 \end{bmatrix},$$

where

$$\begin{aligned} Z_1 &\triangleq P_1 Q_1 + P_{12} Q_{12}^T, & Z_{12} &\triangleq P_1 Q_{12} + P_{12} Q_2, \\ Z_{21} &\triangleq P_{12}^T Q_1 + P_2 Q_{12}^T, & Z_2 &\triangleq P_{12}^T Q_{12} + P_2 Q_2. \end{aligned}$$

Thus, with $\lambda = 1$ the stationarity conditions are given by

$$\frac{\partial \mathcal{L}}{\partial Q} = (\bar{A} + \gamma^{-2}Q\bar{R})^T P + P(\bar{A} + \gamma^{-2}Q\bar{R}) + \bar{R} = 0, \quad (\text{A.4})$$

$$\frac{\partial \mathcal{L}}{\partial A_m} = Z_2 = 0, \quad (\text{A.5})$$

$$\frac{\partial \mathcal{L}}{\partial B_m} = P_{12}^T B V + P_2 B_m V = 0, \quad (\text{A.6})$$

$$\begin{aligned} \frac{\partial \mathcal{L}}{\partial C_m} &= 2RC_m Q_2 + 2\gamma^{-2}RC_m Z_{12}^T Q_{12} - 2RCQ_{12} - \gamma^{-2}RCZ_1^T Q_{12} \\ &\quad - \gamma^{-2}RCQ_1 Z_{12} - \gamma^{-2}RCZ_{21}^T Q_2 = 0. \end{aligned} \quad (\text{A.7})$$

Expanding (2.16) and (A.4) yields

$$0 = AQ_1 + Q_1A^T + \gamma^{-2}(Q_1C^T - Q_{12}C_m^T)R(Q_1C^T - Q_{12}C_m^T)^T + BV_mB_m^T, \quad (A.8)$$

$$0 = AQ_{12} + Q_{12}A_m^T + \gamma^{-2}Q_1C^TRCQ_{12} - \gamma^{-2}Q_{12}C_m^TRCQ_{12} - \gamma^{-2}Q_1C^TRC_mQ_2 \\ + \gamma^{-2}Q_{12}C_m^TRC_mQ_2, \quad (A.9)$$

$$0 = A_mQ_2 + Q_2A_m^T + \gamma^{-2}(Q_{12}^TC^T - Q_2C_m^T)R(Q_{12}^TC^T - Q_2C_m^T)^T + B_mVB_m^T, \quad (A.10)$$

$$0 = A^TP_1 + P_1A + \gamma^{-2}C^TRCZ_1^T - \gamma^{-2}C^TRC_mZ_{12}^T + \gamma^{-2}Z_1C^TRC \\ - \gamma^{-2}Z_{12}C_m^TRC + C^TRC, \quad (A.11)$$

$$0 = A^TP_{12} + P_{12}A_m + \gamma^{-2}C^TRCZ_{21}^T - \gamma^{-2}Z_1C^TRC_m + \gamma^{-2}Z_{12}C_m^TRC_m - C^TRC_m, \quad (A.12)$$

$$0 = A_m^TP_2 + P_2A_m - \gamma^{-2}C_m^TRCZ_{21}^T - \gamma^{-2}Z_{21}C^TRC_m + C_m^TRC_m. \quad (A.13)$$

Now define the $n \times n$ matrices

$$Q \triangleq Q_1 - Q_{12}Q_2^{-1}Q_{12}^T, \quad P \triangleq P_1 - P_{12}P_2^{-1}P_{12}^T,$$

$$\hat{Q} \triangleq Q_{12}Q_2^{-1}Q_{12}^T, \quad \hat{P} \triangleq P_{12}P_2^{-1}P_{12}^T,$$

$$r \triangleq -Q_{12}Q_2^{-1}P_2^{-1}P_{12}^T,$$

and the $n_m \times n$, $n_m \times n_m$ and $n_m \times n$ matrices

$$G \triangleq Q_2^{-1}Q_{12}^T, \quad M \triangleq Q_2P_2, \quad \Gamma \triangleq -P_2^{-1}P_{12}^T.$$

The existence of Q_2^{-1} and P_2^{-1} follows from the fact that (A_m, B_m, C_m) is minimal. See Bernstein and Haddad, 1988, and Hyland and Bernstein, 1985, for details. Note that $r = G^T\Gamma$. Clearly, Q, P, \hat{Q} , and \hat{P} are symmetric and nonnegative definite.

Next note that with the above definitions, (A.5) implies (3.3) and that (3.2) holds. Hence $r = G^T\Gamma$ is idempotent, i.e., $r^2 = r$. Sylvester's inequality yields (3.18). Note also that (3.7) and (3.8) hold.

The components of Q and P can be written in terms of $Q, P, \hat{Q}, \hat{P}, G$, and Γ as

$$Q_1 = Q + \hat{Q}, \quad P_1 = P + \hat{P}, \quad (A.14)$$

$$Q_{12} = \hat{Q}\Gamma^T, \quad P_{12} = -\hat{P}G^T, \quad (A.15)$$

$$Q_2 = \Gamma\hat{Q}\Gamma^T, \quad P_2 = G\hat{P}G^T. \quad (A.16)$$

Next note that by using (A.14)–(A.16), (A.7) becomes

$$C_m \hat{S} = C[I_n + \gamma^{-2}(Q + \hat{Q})P]G^T,$$

where

$$\hat{S} \triangleq I_{n_m} + \gamma^{-2}\Gamma\hat{Q}PG^T.$$

To prove that \hat{S} is invertible use (3.7) and (3.4) and note that

$$\begin{aligned} I_{n_m} + \gamma^{-2}\Gamma\hat{Q}PG^T &= I_{n_m} + \gamma^{-2}\Gamma\hat{Q}\Gamma^T PG^T \\ &= I_{n_m} + \gamma^{-2}(\Gamma\hat{Q}\Gamma^T)(GPG^T). \end{aligned}$$

Since $\Gamma\hat{Q}\Gamma^T$ and GPG^T are nonnegative definite, their product has nonnegative eigenvalues. Thus each eigenvalue of $I_{n_m} + \gamma^{-2}\Gamma\hat{Q}PG^T$ is real and is greater than unity. Hence \hat{S} is invertible. Now note that by using (3.3) and (3.4) it can be shown that

$$G^T \hat{S}^{-1} \Gamma = S\tau.$$

The expressions (3.11), (3.12) and (3.13) follow from (A.6), (A.7), (3.9) and the definition of \mathcal{Q} by using the above identities. Next, computing either Γ (A.9)–(A.10) or G (A.12)–(A.13) yields (3.10). Substituting this expression for A_m into (A.8)–(A.13) it follows that (A.10) = Γ (A.9) and (A.13) = G (A.12). Thus, (A.10) and (A.13) are superfluous and can be omitted. Next, using (A.8) + $G^T \Gamma$ (A.9) G – (A.9) G – $[(A.9)G]^T$ and $G^T \Gamma$ (A.9) G – (A.9) G – $[(A.9)G]^T$ yields (3.14) and (3.16). Using (A.11) + $\Gamma^T G$ (A.12) Γ – (A.12) Γ – $[(A.12)\Gamma]^T$ and $\Gamma^T G$ (A.12) Γ – (A.12) Γ – $[(A.12)\Gamma]^T$ yields (3.15) and (3.17).

Finally, to prove the converse we use (3.10)–(3.18) to obtain (2.16) and (A.4)–(A.7). Let $A_m, B_m, C_m, G, \Gamma, \tau, Q, P, \hat{Q}, \hat{P}, \mathcal{Q}$ be as in the statement of Theorem 3.1 and define $Q_1, Q_{12}, Q_2, P_1, P_{12}, P_2$ by (A.14)–(A.16). Using (3.3), (3.11) and (3.12) it is easy to verify (A.6) and (A.7). Finally, substitute the definitions of $Q, P, \hat{Q}, \hat{P}, G, \Gamma$, and τ into (3.14)–(3.17) along with (3.3), (3.4), (3.7) and (3.8) to obtain (2.16) and (A.4). Finally, note that

$$\mathcal{Q} = \begin{bmatrix} Q & 0_{n \times n_m} \\ 0_{n_m \times n} & 0_{n_m} \end{bmatrix} + \begin{bmatrix} I_n \\ \Gamma \end{bmatrix} \hat{Q} \begin{bmatrix} I_n & \Gamma^T \end{bmatrix},$$

which shows that $\mathcal{Q} \geq 0$. \square

References

- Adamjan, M., Arov, D. Z., and Krein, M. G., 1971, *Math. USSR Sbornik*, **15**, 31; 1978, *Amer. Math. Soc. Transl.*, **111**, 133.
- Anderson, B. D. O., 1986, *Sys. Contr. Lett.*, **7**, 247.
- Ball, J. A., and Ran, C. M., 1987, *SIAM J. Contr. Optim.*, **25**, 362.
- Bernstein, D. S., and Haddad, W. M., 1988, "LQG Control with an H_∞ Performance Bound: A Riccati Equation Approach," *IEEE Trans. Autom. Contr.*, submitted.
- Brockett, R. W., 1970, *Finite Dimensional Linear Systems*, (Wiley).
- Glover, K., 1984, *Int. J. Contr.*, **39**, 1115.
- Hung, Y. S., and Glover, K., 1986, *Sys. Contr. Lett.*, **7**, 165.
- Hyland, D. C., and Bernstein, D. S., 1985, *IEEE Trans. Autom. Contr.*, **30**, 1201.
- Khargonekar, P. P., Petersen, I. R., and Zhou, K., 1987, "Robust Stabilization of Uncertain Systems and H^∞ Optimal Control," preprint.
- Kung, S.-Y., and Lin, D. W., 1981, *IEEE Trans. Autom. Contr.*, **26**, 832.
- Latham, G. A., and Anderson, B. D. O., 1985, *Sys. Contr. Lett.*, **5**, 229.
- Lin, D. W. and Kung, S.-Y., 1982, *Cir. Sys. Sig. Proc.*, **1**, 407.
- Parker, P. J., and Anderson, B. D. O., 1987, *Int. J. Contr.*, **46**, 1783.
- Petersen, I. R., 1987, *IEEE Trans. Autom. Contr.*, **32**, 427.
- Rao, C. R., and Mitra, S. K., 1971, *Generalized Inverse of Matrices and Its Applications*, (John Wiley and Sons).
- Richter, S., 1987, *Proc. Amer. Contr. Conf.*, June 1987, Minneapolis, MN, 1527.
- Willems, J. C., 1971, *IEEE Trans. Autom. Contr.*, **16**, 621.
- Wilson, D. A., 1970, *Proc. IEE*, **117**, 1161.
- Wilson, D. A., 1985, *Int. J. Contr.*, **42**, 65.
- Wilson, D. A., 1988, "System Order Reduction Using the Hilbert-Schmidt Norm of the Hankel Operator," preprint.
- Wilson, D. A., 1988a, "Convolution and Hankel Operator Norms for Linear Systems," *IEEE Trans. Autom. Contr.*, to appear.
- Wonham, W. M., 1979, *Linear Multivariable Control: A Geometric Approach*, (Springer-Verlag)

March 1988
Revised June 1988

Steady-State Kalman Filtering with an H_∞ Error Bound

by

Dennis S. Bernstein
Harris Corporation
Government Aerospace Systems Division
MS 22/4848
Melbourne, FL 32902

Wassim M. Haddad
Department of Mechanical and
Aerospace Engineering
Florida Institute of Technology
Melbourne, FL 32901

Abstract

An estimator design problem is considered which involves both L_2 (least squares) and H_∞ (worst-case frequency-domain) aspects. Specifically, the goal of the problem is to minimize an L_2 state-estimation error criterion subject to a prespecified H_∞ constraint on the state-estimation error. The H_∞ estimation-error constraint is embedded within the optimization process by replacing the covariance Lyapunov equation by a Riccati equation whose solution leads to an upper bound on the L_2 state-estimation error. The principal result is a sufficient condition for characterizing fixed-order (i.e., full- and reduced-order) estimators with bounded L_2 and H_∞ estimation error. The sufficient condition involves a system of modified Riccati equations coupled by an oblique projection, i.e., idempotent matrix. When the H_∞ constraint is absent, the sufficient condition specializes to the L_2 state-estimation result given in [2].

Keywords: Kalman filter, H_∞ norm, reduced-order state estimation, optimal projection equations, Hankel norm.

Supported in part by the Air Force Office of Scientific Research under contract F49620-86-C-0002.

1. Introduction

One of the fundamental problems in dynamic systems theory is the observation of state variables. Although an extensive theoretical foundation has been developed for the quadratic (least squares) error criterion, state estimation with a worst-case frequency-domain design objective has apparently not been considered. In the present paper we thus extend the least squares formulation to include a frequency-domain bound on the state-estimation error. The underlying idea involves the application of state-space techniques which have recently been developed for H_∞ control design in [1,4-6]. The results of the present paper are thus complementary to the results obtained in [1].

The principal result of the present paper is a sufficient condition which yields full- and reduced-order estimators satisfying an optimized L_2 error bound as well as a prespecified H_∞ error bound. In the full-order case, the H_∞ -constrained estimator involves a modified Riccati equation which specializes to the standard steady-state Kalman filter when the H_∞ constraint is absent. In the reduced-order case the H_∞ -constrained result leads to a direct generalization of the optimal projection approach developed in [2] for the unconstrained L_2 state-estimation problem. While the L_2 -optimal reduced-order state estimator was characterized in [2] by means of a coupled system of one modified Riccati equation and two modified Lyapunov equations, the H_∞ -constrained solution involves a coupled system consisting of three modified Riccati equations and one modified Lyapunov equation. As in [2], the coupling is due to the presence of an oblique projection (idempotent matrix) with additional coupling now arising from the H_∞ constraint. When the H_∞ constraint is sufficiently relaxed, these conditions again specialize directly to those given in [2].

We note that the development in the present paper is limited to the case in which the plant is asymptotically stable. These results can also be extended to the unstable plant case, although with additional complexity. This case will thus be treated in a future paper.

The contents of the paper are as follows. After collecting notation in Section 2, the statement of the H_∞ -Constrained State-Estimation Problem is given in Section 3. The principal result of this section (Lemma 3.1) shows that if the algebraic Lyapunov equation for the covariance is replaced by a modified Riccati equation possessing a nonnegative-definite solution, then the H_∞ estimation-error constraint is enforced and the L_2 state-estimation error criterion is bounded above by an auxiliary cost function. The problem of determining a reduced-order estimator which minimizes this upper bound subject to the Riccati equation constraint is considered in Section 4 as the Auxiliary Minimization Problem. Necessary conditions for the Auxiliary Minimization Problem

(Theorem 4.1) are given in the form of a coupled system of modified algebraic Riccati equations. To develop connections with standard Kalman filter theory the full-order estimator result is also given. In Section 5 the necessary conditions of Theorem 4.1 are combined with Lemma 3.1 to yield sufficient conditions for bounded H_∞ and L_2 estimation error. Although our result gives sufficient conditions for H_∞ estimation error, we also state hypotheses under which these conditions are also necessary (Proposition 5.1).

2. Notation and Definitions

| | |
|---|---|
| $\mathbb{R}, \mathbb{R}^{r \times s}, \mathbb{R}^r, \mathbb{E}$ | real numbers, $r \times s$ real matrices, $\mathbb{R}^{r \times 1}$, expected value |
| $I_r, (\)^T, 0_{r \times s}, 0_r$ | $r \times r$ identity matrix, transpose, $r \times s$ zero matrix, $0_{r \times r}$ |
| tr | trace |
| $\sigma_{\max}(Z)$ | largest singular value of matrix Z |
| $\lambda_{\max}(Z)$ | largest eigenvalue of matrix Z with real spectrum |
| $\ Z\ _F$ | $[\text{tr } ZZ^T]^{\frac{1}{2}}$ (Frobenius matrix norm) |
| $\ H(s)\ _\infty$ | $\sup_{\omega \in \mathbb{R}} \sigma_{\max}[H(j\omega)]$ |
| $\mathbb{S}^r, \mathbb{N}^r, \mathbb{P}^r$ | $r \times r$ symmetric, nonnegative-definite, positive-definite matrices |
| $Z_1 \leq Z_2, Z_1 < Z_2$ | $Z_2 - Z_1 \in \mathbb{N}^r, Z_2 - Z_1 \in \mathbb{P}^r, Z_1, Z_2 \in \mathbb{S}^r$ |
| $n, \ell, n_e, p, q, r; \tilde{n}$ | positive integers; $n + n_e; n_e \leq n$ |
| $x, y, y_e, x_e, \tilde{x}$ | $n, \ell, q, n_e, \tilde{n}$ -dimensional vectors |
| \tilde{x} | $\begin{bmatrix} x \\ x_e \end{bmatrix}$ |
| A, C | $n \times n, \ell \times n$ matrices |
| D_1, D_2, E | $n \times p, \ell \times p, r \times q$ matrices |
| L | $q \times n$ matrix |
| A_e, B_e, C_e | $n_e \times n_e, n_e \times \ell, q \times n_e$ matrices |
| \tilde{A} | $\begin{bmatrix} A & 0_{n \times n_e} \\ B_e C & A_e \end{bmatrix}$ |
| \tilde{D}, \tilde{E} | $\begin{bmatrix} D_1 \\ B_e D_2 \end{bmatrix}, [EL \quad -EC_e]$ |
| R | $E^T E$, estimation error weighting in \mathbb{P}^q |
| $w(\cdot)$ | p -dimensional standard white noise process |
| V_1, V_2 | intensity of $D_1 w(\cdot), D_2 w(\cdot)$; $V_1 = D_1 D_1^T \in \mathbb{P}^n, V_2 = D_2 D_2^T \in \mathbb{P}^\ell$ |

| | |
|------------------------|--|
| V_{12} | cross intensity of $D_1 w(\cdot), D_2 w(\cdot)$; $V_{12} = D_1 D_2^T \in \mathbb{R}^{n \times \ell}$ |
| \tilde{R}, \tilde{V} | $\begin{bmatrix} L^T R L & -L^T R C_e \\ -C_e^T R L & C_e^T R C_e \end{bmatrix}, \quad \begin{bmatrix} V_1 & V_{12} B_e^T \\ B_e V_{12}^T & B_e V_2 B_e^T \end{bmatrix}$ |
| γ | positive constant |

3. Statement of the Problem

In this section we introduce the reduced-order state-estimation problem with a constraint on the H_∞ norm of the state-estimation error. Specifically, the transfer function between disturbances and error states is constrained to have H_∞ norm less than γ . In this paper it is assumed that the plant is asymptotically stable, i.e., the eigenvalues of A are in the open left half plane.

H_∞ -Constrained State-Estimation Problem. Given the n th-order observed system

$$\dot{x}(t) = Ax(t) + D_1 w(t), \quad (3.1)$$

$$y(t) = Cx(t) + D_2 w(t), \quad (3.2)$$

where $t \in [0, \infty)$, determine an n_e th-order state estimator

$$\dot{x}_e(t) = A_e x_e(t) + B_e y(t), \quad (3.3)$$

$$y_e(t) = C_e x_e(t), \quad (3.4)$$

where $n_e \leq n$, which satisfies the following design criteria:

- (i) A_e is asymptotically stable;
- (ii) the $r \times p$ transfer function

$$H(s) \triangleq \tilde{E}(sI_{\tilde{n}} - \tilde{A})^{-1} \tilde{D} \quad (3.5)$$

from disturbances $w(t)$ to error states $E[Lx(t) - y_e(t)] = \tilde{E}\tilde{x}(t)$ satisfies the constraint.

$$\|H(s)\|_\infty \leq \gamma, \quad (3.6)$$

where $\gamma > 0$ is a given constant; and

- (iii) the L_2 state-estimation error criterion

$$J(A_e, B_e, C_e) \triangleq \lim_{t \rightarrow \infty} \mathbb{E} \left\{ [Lx(t) - y_e(t)]^T R [Lx(t) - y_e(t)] \right\} \quad (3.7)$$

is minimized.

It is useful to note that the augmented system (3.1)–(3.4) can be written as

$$\dot{\tilde{x}}(t) = \tilde{A}\tilde{x}(t) + \tilde{D}w(t), \quad t \in [0, \infty), \quad (3.8)$$

and that (3.7) is equivalent to

$$J(A_e, B_e, C_e) = \lim_{t \rightarrow \infty} \mathbb{E} \left\{ [\tilde{E}\tilde{x}(t)]^T [\tilde{E}\tilde{x}(t)] \right\} = \lim_{t \rightarrow \infty} \mathbb{E} [\tilde{x}^T(t) \tilde{R}\tilde{x}(t)]. \quad (3.9)$$

Furthermore, if A_e is asymptotically stable for a given estimator (A_e, B_e, C_e) then the L_2 state-estimation error criterion is given by

$$J(A_e, B_e, C_e) = \text{tr } \tilde{Q}\tilde{R}, \quad (3.10)$$

where the steady-state covariance defined by

$$\tilde{Q} \triangleq \lim_{t \rightarrow \infty} \mathbb{E} [\tilde{x}(t)\tilde{x}^T(t)] \quad (3.11)$$

satisfies the $\tilde{n} \times \tilde{n}$ Lyapunov equation

$$0 = \tilde{A}\tilde{Q} + \tilde{Q}\tilde{A}^T + \tilde{V}. \quad (3.12)$$

Using (3.10) and (3.12) we now show that the criterion (3.7) is an error measure involving the impulse response of (3.8) with respect to an L_2 norm.

Proposition 3.1. If A_e is asymptotically stable then the L_2 state-estimation error criterion (3.7) can be written as

$$J(A_e, B_e, C_e) = \int_0^\infty \|\tilde{E}e^{\tilde{A}t}\tilde{D}\|_F^2 dt. \quad (3.13)$$

Proof. It need only be noted that (3.10) is equivalent to

$$\text{tr} \int_0^\infty e^{\tilde{A}t}\tilde{V}e^{\tilde{A}^T t} dt \tilde{R} = \text{tr} \int_0^\infty (\tilde{E}e^{\tilde{A}t}\tilde{D})(\tilde{E}e^{\tilde{A}t}\tilde{D})^T dt,$$

which is equivalent to (3.13). \square

The key step in enforcing (3.6) is to replace the algebraic Lyapunov equation (3.12) by an algebraic Riccati equation. Justification for this technique is provided by the following result.

Lemma 3.1. Let (A_e, B_e, C_e) be given and assume there exists $Q \in \mathbb{R}^{\tilde{n} \times \tilde{n}}$ satisfying

$$Q \in \mathbb{N}^{\tilde{n}} \quad (3.14)$$

and

$$0 = \tilde{A}Q + Q\tilde{A}^T + \gamma^{-2}Q\tilde{R}Q + \tilde{V}. \quad (3.15)$$

Then

$$(\tilde{A}, [\gamma^{-2}Q\tilde{R}Q + \tilde{V}]^{\frac{1}{2}}) \text{ is stabilizable} \quad (3.16)$$

if and only if

$$A_e \text{ is asymptotically stable.} \quad (3.17)$$

Furthermore, in this case

$$\|H(s)\|_{\infty} \leq \gamma, \quad (3.18)$$

$$\tilde{Q} \leq Q, \quad (3.19)$$

and

$$J(A_e, B_e, C_e) \leq J(A_e, B_e, C_e, Q), \quad (3.20)$$

where

$$J(A_e, B_e, C_e, Q) \triangleq \text{tr } Q\tilde{R}. \quad (3.21)$$

Proof. Using the assumed existence a nonnegative-definite solution to (3.15) and the stabilizability condition (3.16), it follows from the dual of Lemma 12.2 of [9] that \tilde{A} is asymptotically stable. Since \tilde{A} is lower block triangular, \tilde{A} asymptotically stable implies A_e is asymptotically stable. Conversely, since A is assumed to be asymptotically stable, (3.17) implies \tilde{A} is asymptotically stable and thus (3.16) holds. The proof of (3.18) follows from a standard manipulation of (3.15); for details see Lemma 1 of [8]. To prove (3.19) subtract (3.12) from (3.15) to obtain

$$0 = \tilde{A}(Q - \tilde{Q}) + (Q - \tilde{Q})\tilde{A}^T + \gamma^{-2}Q\tilde{R}Q, \quad (3.22)$$

which, since \tilde{A} is asymptotically stable, is equivalent to

$$Q - \tilde{Q} = \int_0^{\infty} e^{\tilde{A}t} [\gamma^{-2}Q\tilde{R}Q] e^{\tilde{A}^T t} dt \geq 0. \quad (3.23)$$

Finally, (3.20) follows immediately from (3.19). \square

Lemma 3.1 shows that the H_∞ constraint is automatically enforced when a nonnegative-definite solution to (3.15) can be shown to exist. Furthermore, the solution \mathcal{Q} provides an upper bound for the steady-state covariance $\tilde{\mathcal{Q}}$ along with a bound on the L_2 state-estimation error criterion. Next, we present a partial converse of Lemma 3.1 which guarantees the existence of a nonnegative-definite solution to (3.15) when (3.18) is satisfied.

Lemma 3.2. Let (A_e, B_e, C_e) be given, suppose A_e is asymptotically stable, and assume the H_∞ state-estimation error constraint (3.18) is satisfied. Then there exists a unique nonnegative-definite solution \mathcal{Q} satisfying (3.15). Furthermore, $(\tilde{A} + \gamma^{-2}\mathcal{Q}\tilde{R}, \tilde{D})$ is stabilizable if and only if $\tilde{A} + \gamma^{-2}\mathcal{Q}\tilde{R}$ is asymptotically stable.

Proof. The result is an immediate consequence of Theorems 3 and 2 of [3], pp. 150 and 167, along with the dual of Lemma 12.2 of [9]. \square

Finally, we show that the quadratic term $\gamma^{-2}\mathcal{Q}\tilde{R}\mathcal{Q}$ in (3.15) also constrains the Hankel norm of the estimation error $E[Lx(t) - y_e(t)]$ when \mathcal{Q} is positive definite. To show this let $\tilde{P} \in \mathbb{R}^{\tilde{n}}$ be the observability Gramian for the augmented system $(\tilde{A}, \tilde{D}, \tilde{E})$ which satisfies

$$0 = \tilde{A}^T \tilde{P} + \tilde{P} \tilde{A} + \tilde{R}. \quad (3.24)$$

Proposition 3.2. Let (A_e, B_e, C_e) be given and assume there exists $\mathcal{Q} \in \mathbb{P}^{\tilde{n}}$ satisfying (3.15) and (3.16) or, equivalently, (3.17). Then

$$\lambda_{\max}^{\frac{1}{2}}(\tilde{P}\tilde{\mathcal{Q}}) \leq \gamma. \quad (3.25)$$

Proof. Since \mathcal{Q} is invertible, (3.15) implies

$$0 = \gamma^2 \tilde{A}^T \mathcal{Q}^{-1} + \gamma^2 \mathcal{Q}^{-1} \tilde{A} + \gamma^2 \mathcal{Q}^{-1} \tilde{V} \mathcal{Q}^{-1} + \tilde{R}. \quad (3.26)$$

Next, subtract (3.24) from (3.26) to obtain

$$0 = \tilde{A}^T (\gamma^2 \mathcal{Q}^{-1} - \tilde{P}) + (\gamma^2 \mathcal{Q}^{-1} - \tilde{P}) \tilde{A} + \gamma^2 \mathcal{Q}^{-1} \tilde{V} \mathcal{Q}^{-1}, \quad (3.27)$$

which, since \tilde{A} is asymptotically stable, is equivalent to

$$\gamma^2 \mathcal{Q}^{-1} - \tilde{P} = \int_0^\infty e^{\tilde{A}^T t} [\gamma^2 \mathcal{Q}^{-1} \tilde{V} \mathcal{Q}^{-1}] e^{\tilde{A} t} dt \geq 0. \quad (3.28)$$

Thus (3.28) implies $\tilde{P} \leq \gamma^2 Q^{-1}$ or, equivalently, $Q^{\frac{1}{2}} \tilde{P} Q^{\frac{1}{2}} \leq \gamma^2 I_n$. Hence,

$$\gamma^2 \geq \lambda_{\max}(Q^{\frac{1}{2}} \tilde{P} Q^{\frac{1}{2}}) = \lambda_{\max}(\tilde{P}^{\frac{1}{2}} Q \tilde{P}^{\frac{1}{2}}) \geq \lambda_{\max}(\tilde{P}^{\frac{1}{2}} \tilde{Q} \tilde{P}^{\frac{1}{2}}) = \lambda_{\max}(\tilde{P} \tilde{Q}). \quad \square$$

4. The Auxiliary Minimization Problem and Necessary Conditions for Optimality

As discussed in the previous section, the replacement of (3.12) by (3.15) enforces the H_∞ state-estimation error constraint and results in an upper bound for the L_2 state-estimation error criterion. That is, given an estimator (A_e, B_e, C_e) satisfying the H_∞ estimation constraint, the *actual* L_2 state-estimation error criterion is guaranteed to be no worse than the bound given by $J(A_e, B_e, C_e, Q)$ if (3.15) is solvable. Hence, $J(A_e, B_e, C_e, Q)$ can be interpreted as an *auxiliary cost* which leads to the following optimization problem.

Auxiliary Minimization Problem. Determine (A_e, B_e, C_e, Q) which minimizes $J(A_e, B_e, C_e, Q)$ subject to (3.14) and (3.15).

It follows from Lemma 3.1 that the satisfaction of (3.14)–(3.16) leads to 1) A_e stable; 2) H_∞ estimation error bound γ ; and 3) an upper bound (3.21) for the L_2 state-estimation error criterion. Hence it remains to determine (A_e, B_e, C_e) which minimizes $J(A_e, B_e, C_e, Q)$ and thus provides an optimized bound for the *actual* L_2 criterion $J(A_e, B_e, C_e)$. Rigorous derivation of the necessary conditions for the Auxiliary Minimization Problem requires additional technical assumptions. Specifically, we restrict (A_e, B_e, C_e, Q) to the open set

$$S \triangleq \{(A_e, B_e, C_e, Q) : Q \in \mathbb{P}^n, \tilde{A} + \gamma^{-2} Q \tilde{R} \text{ is asymptotically stable, and } (A_e, B_e, C_e) \text{ is controllable and observable}\}.$$

Remark 4.1. The set S constitutes sufficient conditions under which the Lagrange multiplier technique is applicable to the Auxiliary Minimization Problem. Specifically, the requirement that Q be positive definite replaces (3.14) by an open set constraint, the stability of $\tilde{A} + \gamma^{-2} Q \tilde{R}$ serves as a normality condition, and (A_e, B_e, C_e) minimal is a nondegeneracy condition.

The following Lemma is needed for the statement of the main result.

Lemma 4.1. Let $\hat{Q}, \hat{P} \in \mathbb{R}^n$ and suppose $\text{rank } \hat{Q} \hat{P} = n_e$. Then there exist $n_e \times n$ G, Γ and $n_e \times n_e$ invertible M , unique except for a change of basis in \mathbb{R}^{n_e} , such that

$$\hat{Q} \hat{P} = G^T M \Gamma, \quad (4.1)$$

$$\Gamma G^T = I_{n_e}. \quad (4.2)$$

Furthermore, the $n \times n$ matrices

$$\tau \triangleq G^T \Gamma, \quad (4.3)$$

$$\tau_\perp \triangleq I_n - \tau, \quad (4.4)$$

are idempotent and have rank n_e and $n - n_e$, respectively. If, in addition,

$$\text{rank } \hat{Q} = \text{rank } \hat{P} = n_e, \quad (4.5)$$

then

$$\hat{Q} = \tau \hat{Q}, \quad \hat{P} = \hat{P} \tau. \quad (4.6), (4.7)$$

Finally, if $P \in \mathbb{R}^n$ then the inverse

$$S \triangleq (I_n + \gamma^{-2} \hat{Q} P)^{-1} \quad (4.8)$$

exists.

Proof. Conditions (4.1)–(4.7) are a direct consequence of Theorem 6.2.5 of [7]. To prove that the inverse in (4.8) exists, note that since the eigenvalues of $\hat{Q} P$ coincide with the eigenvalues of the nonnegative-definite matrix $P^{\frac{1}{2}} \hat{Q} P^{\frac{1}{2}}$, it follows that $\hat{Q} P$ has nonnegative eigenvalues. Thus, the eigenvalues of $I_n + \gamma^{-2} \hat{Q} P$ are all greater than one so that the above inverse exists. \square

Finally, for arbitrary $Q \in \mathbb{R}^{n \times n}$ define the notation

$$Q_* \triangleq Q C^T + V_{12}, \quad \Sigma \triangleq L^T R L. \quad (4.9)$$

Theorem 4.1. If $(A_*, B_*, C_*, Q) \in S$ solves the Auxiliary Minimization Problem then there exist $Q, P, \hat{Q}, \hat{P} \in \mathbb{R}^n$ such that

$$A_* = \Gamma(A - Q_* V_2^{-1} C - \gamma^{-4} Q \Sigma Q P S) G^T, \quad (4.10)$$

$$B_* = \Gamma Q_* V_2^{-1}, \quad (4.11)$$

$$C_* = L(I_n + \gamma^{-2} Q P S) G^T, \quad (4.12)$$

$$Q = \begin{bmatrix} Q + \hat{Q} & \hat{Q} \Gamma^T \\ \Gamma \hat{Q} & \Gamma \hat{Q} \Gamma^T \end{bmatrix}, \quad (4.13)$$

and such that Q, P, \hat{Q}, \hat{P} satisfy

$$0 = AQ + QA^T + V_1 + \gamma^{-2}Q\Sigma Q - Q_\alpha V_2^{-1}Q_\alpha^T + \tau_\perp Q_\alpha V_2^{-1}Q_\alpha^T \tau_\perp^T, \quad (4.14)$$

$$0 = A^T P + PA - \gamma^{-4}S^T P Q \Sigma Q P S + \tau_\perp^T (I_n + \gamma^{-2}Q P S)^T \Sigma (I_n + \gamma^{-2}Q P S) \tau_\perp, \quad (4.15)$$

$$0 = (A - \gamma^{-4}Q \Sigma Q P S) \hat{Q} + \hat{Q} (A - \gamma^{-4}Q \Sigma Q P S)^T + \gamma^{-4} \hat{Q} S^T P Q \Sigma Q P S \hat{Q} \\ + Q_\alpha V_2^{-1} Q_\alpha^T - \tau_\perp Q_\alpha V_2^{-1} Q_\alpha^T \tau_\perp^T, \quad (4.16)$$

$$0 = (A - Q_\alpha V_2^{-1} C + \gamma^{-2}Q \Sigma)^T \hat{P} + \hat{P} (A - Q_\alpha V_2^{-1} C + \gamma^{-2}Q \Sigma) \\ + (I_n + \gamma^{-2}Q P S)^T \Sigma (I_n + \gamma^{-2}Q P S) - \tau_\perp^T (I_n + \gamma^{-2}Q P S)^T \Sigma (I_n + \gamma^{-2}Q P S) \tau_\perp, \quad (4.17)$$

$$\text{rank } \hat{Q} = \text{rank } \hat{P} = \text{rank } \hat{Q} \hat{P} = n_e. \quad (4.18)$$

Furthermore, the auxiliary cost is given by

$$J(A_e, B_e, C_e, Q) = \text{tr } L^T R L (Q + \gamma^{-4}Q P S \hat{Q} S^T P Q). \quad (4.19)$$

Conversely, if there exist $Q, P, \hat{Q}, \hat{P} \in \mathbb{R}^{n \times n}$ satisfying (4.14)–(4.18), then (A_e, B_e, C_e, Q) given by (4.10)–(4.13) satisfies (3.14) and (3.15) with the auxiliary cost (3.21) given by (4.19).

Proof. See Appendix A. \square

Remark 4.2. Theorem 4.1 presents necessary conditions for the Auxiliary Minimization Problem which explicitly synthesize extremal full- and reduced-order estimators (A_e, B_e, C_e) . If the H_∞ estimation constraint is sufficiently relaxed, i.e., $\gamma \rightarrow \infty$, then $S = I_n$. In this case equations (4.16) and (4.17) become decoupled from (4.15) and thus (4.15) becomes superfluous. Furthermore, (4.14), (4.16) and (4.17) specialize to the optimal projection equations obtained in [2].

As discussed in [2], in the full-order (Kalman Filter) case $n_e = n$, $G = \Gamma^{-1}$ and thus $G = \Gamma = \tau = I_n$ and $\tau_\perp = 0$ without loss of generality. To develop further connections with the standard Kalman filter theory assume

$$V_{12} = 0. \quad (4.20)$$

In this case (4.15) implies that $P = 0$ so that the gain expressions (4.10)–(4.12) become

$$A_e = A - Q C^T V_2^{-1} C, \quad (4.21)$$

$$B_e = Q C^T V_2^{-1}, \quad (4.22)$$

$$C_e = L, \quad (4.23)$$

while equations (4.14)–(4.16) and auxiliary cost (4.19) specialize to

$$0 = AQ + QA^T + V_1 + \gamma^{-2}QL^T RLQ - QC^T V_2^{-1} CQ, \quad (4.24)$$

$$J(A_e, B_e, C_e, Q) = \text{tr } L^T RLQ. \quad (4.25)$$

Remark 4.3. Note that the necessary conditions for the full-order problem involve one modified Riccati equation. This equation is similar to the observer Riccati equation with the additional quadratic term $\gamma^{-2}QL^T RLQ$. Finally, note that when the H_∞ estimation constraint is sufficiently relaxed, i.e., $\gamma \rightarrow \infty$, (4.24) reduces to the standard observer Riccati equation of steady-state Kalman filter theory.

5. Sufficient Conditions for Combined L_2/H_∞ Estimation

In this section we combine Lemma 3.1 with the converse of Theorem 4.1 to obtain our main result guaranteeing constrained H_∞ state-estimation error and an optimized L_2 state-estimation error bound.

Theorem 5.1. Suppose there exist $Q, P, \hat{Q}, \hat{P} \in \mathbb{R}^n$ satisfying (4.14)–(4.18) and let (A_e, B_e, C_e, Q) be given by (4.10)–(4.13). Then $(\tilde{A}, [\gamma^{-2}Q\tilde{R}Q + \tilde{V}]^{\frac{1}{2}})$ is stabilizable if and only if A_e is asymptotically stable. In this case, the transfer function $H(s)$ defined by (3.5) satisfies the H_∞ state-estimation error constraint

$$\|H(s)\|_\infty \leq \gamma, \quad (5.1)$$

and the L_2 state-estimation error criterion (3.7) satisfies the bound

$$J(A_e, B_e, C_e) \leq \text{tr } L^T RL(Q + \gamma^{-4}QPS\hat{Q}S^T P Q). \quad (5.2)$$

Proof. The converse portion of Theorem 4.1 implies that Q given by (4.13) satisfies (3.14) and (3.15). It now follows from Lemma 3.1 that the stabilizability condition (3.16) is equivalent to the asymptotic stability of A_e , the H_∞ state-estimation error constraint (3.18) holds, and the L_2 state-estimation error criterion (3.7) satisfies the bound (3.20) which, by (4.19), is equivalent to (5.2). \square

In applying Theorem 5.1 the principal issue concerns conditions on the problem data under which the coupled Riccati equations (4.14)–(4.17) possess nonnegative-definite solutions. Clearly,

for γ sufficiently large, (4.14)–(4.17) approximate the pure least squares problem considered in [2]. The important case of interest, however, involves small γ so that significant H_∞ estimation is enforced. Thus, if (5.1) can be satisfied for a given $\gamma > 0$, it is of interest to know whether one such fixed-order estimator can be obtained by solving (4.14)–(4.17). Lemma 3.2 guarantees that (3.15) possesses a solution for any fixed-order estimator satisfying (5.1). Thus our sufficient conditions will also be *necessary* so long as the Auxiliary Minimization Problem possesses at least one extremal over S . When this is the case we have the following result.

Proposition 5.1. Let γ^* denote the infimum of $\|H(s)\|_\infty$ over all asymptotically stable fixed-order estimators and suppose that the Auxiliary Minimization Problem has an extremal for all $\gamma > \gamma^*$. Then for all $\gamma > \gamma^*$ there exist $Q, P, \hat{Q}, \hat{P} \in \mathbb{R}^n$ satisfying (4.14)–(4.17).

Appendix A: Proof of Theorem 4.1

To optimize (3.21) over the open set S subject to the constraint (3.15), form the Lagrangian

$$\mathcal{L}(A_e, B_e, C_e, Q, P, \lambda) \triangleq \text{tr} \{ \lambda Q \tilde{R} + [\tilde{A}Q + Q\tilde{A}^T + \gamma^{-2} Q \tilde{R} Q + \tilde{V}]P \}, \quad (\text{A.1})$$

where the Lagrange multipliers $\lambda \geq 0$ and $P \in \mathbb{R}^{\tilde{n} \times \tilde{n}}$ are not both zero. We thus obtain

$$\frac{\partial \mathcal{L}}{\partial Q} = (\tilde{A} + \gamma^{-2} Q \tilde{R})^T P + P(\tilde{A} + \gamma^{-2} Q \tilde{R}) + \lambda \tilde{R}. \quad (\text{A.2})$$

Setting $\frac{\partial \mathcal{L}}{\partial Q} = 0$ yields

$$0 = (\tilde{A} + \gamma^{-2} Q \tilde{R})^T P + P(\tilde{A} + \gamma^{-2} Q \tilde{R}) + \lambda \tilde{R}. \quad (\text{A.3})$$

Since $\tilde{A} + \gamma^{-2} Q \tilde{R}$ is assumed to be stable, $\lambda = 0$ implies $P = 0$. Hence, it can be assumed without loss of generality that $\lambda = 1$. Furthermore, P is nonnegative definite.

Now partition $\tilde{n} \times \tilde{n}$ Q, P into $n \times n, n \times n_e$, and $n_e \times n_e$ subblocks as

$$Q = \begin{bmatrix} Q_1 & Q_{12} \\ Q_{12}^T & Q_2 \end{bmatrix}, \quad P = \begin{bmatrix} P_1 & P_{12} \\ P_{12}^T & P_2 \end{bmatrix},$$

and for notational convenience define

$$PQ = \begin{bmatrix} Z_1 & Z_{12} \\ Z_{21} & Z_2 \end{bmatrix},$$

where

$$\begin{aligned} Z_1 &\triangleq P_1 Q_1 + P_{12} Q_{12}^T, & Z_{12} &\triangleq P_1 Q_{12} + P_{12} Q_2, \\ Z_{21} &\triangleq P_{12}^T Q_1 + P_2 Q_{12}^T, & Z_2 &\triangleq P_{12}^T Q_{12} + P_2 Q_2. \end{aligned}$$

Thus, with $\lambda = 1$ the stationarity conditions are given by

$$\frac{\partial \mathcal{L}}{\partial \tilde{Q}} = (\tilde{A} + \gamma^{-2} \tilde{Q} \tilde{R})^T \tilde{P} + \tilde{P}(\tilde{A} + \gamma^{-2} \tilde{Q} \tilde{R}) + \tilde{R} = 0, \quad (\text{A.4})$$

$$\frac{\partial \mathcal{L}}{\partial A_e} = Z_2 = 0, \quad (\text{A.5})$$

$$\frac{\partial \mathcal{L}}{\partial B_e} = Z_{21} C^T + P_{12}^T V_{12} + P_2 B_e V_2 = 0, \quad (\text{A.6})$$

$$\begin{aligned} \frac{\partial \mathcal{L}}{\partial C_e} = & 2RC_e Q_2 + 2\gamma^{-2} RC_e Z_{12}^T Q_{12} - 2RLQ_{12} - \gamma^{-2} RLZ_1^T Q_{12} \\ & - \gamma^{-2} RLQ_1 Z_{12} - \gamma^{-2} RLZ_{21}^T Q_2 = 0. \end{aligned} \quad (\text{A.7})$$

Expanding (3.15) and (A.4) yields

$$0 = AQ_1 + Q_1 A^T + V_1 + \gamma^{-2}(Q_1 L^T - Q_{12} C_e^T)R(Q_1 L^T - Q_{12} C_e^T)^T, \quad (\text{A.8})$$

$$\begin{aligned} 0 = & AQ_{12} + Q_{12} A_e^T + Q_1 C^T B_e^T + V_{12} B_e^T + \gamma^{-2} Q_1 L^T RLQ_{12} - \gamma^{-2} Q_{12} C_e^T RLQ_{12} \\ & - \gamma^{-2} Q_1 L^T RC_e Q_2 + \gamma^{-2} Q_{12} C_e^T RC_e Q_2, \end{aligned} \quad (\text{A.9})$$

$$\begin{aligned} 0 = & A_e Q_2 + Q_2 A_e^T + B_e C Q_{12} + Q_{12}^T C^T B_e^T + B_e V_2 B_e^T \\ & + \gamma^{-2}(Q_{12}^T L^T - Q_2 C_e^T)R(Q_{12}^T L^T - Q_2 C_e^T)^T, \end{aligned} \quad (\text{A.10})$$

$$\begin{aligned} 0 = & A^T P_1 + P_1 A + C^T B_e^T P_{12}^T + P_{12} B_e C + \gamma^{-2} L^T RLZ_1^T + \gamma^{-2} Z_1 L^T RL \\ & - \gamma^{-2} L^T RC_e Z_{12}^T - \gamma^{-2} Z_{12} C_e^T RL + L^T RL, \end{aligned} \quad (\text{A.11})$$

$$\begin{aligned} 0 = & A^T P_{12} + P_{12} A_e + C^T B_e^T P_2 + \gamma^{-2} L^T RLZ_{21}^T - \gamma^{-2} Z_1 L^T RC_e \\ & + \gamma^{-2} Z_{12} C_e^T RC_e - L^T RC_e, \end{aligned} \quad (\text{A.12})$$

$$0 = A_e^T P_2 + P_2 A_e - \gamma^{-2} C_e^T RLZ_{21}^T - \gamma^{-2} Z_{21} L^T RC_e + C_e^T RC_e. \quad (\text{A.13})$$

Now define the $n \times n$ matrices

$$Q \triangleq Q_1 - Q_{12} Q_2^{-1} Q_{12}^T, \quad P \triangleq P_1 - P_{12} P_2^{-1} P_{12}^T,$$

$$\hat{Q} \triangleq Q_{12} Q_2^{-1} Q_{12}^T, \quad \hat{P} \triangleq P_{12} P_2^{-1} P_{12}^T,$$

$$r \triangleq -Q_{12} Q_2^{-1} P_2^{-1} P_{12}^T,$$

and the $n_e \times n$, $n_e \times n_e$, and $n_e \times n$ matrices

$$G \triangleq Q_2^{-1} Q_{12}^T, \quad M \triangleq Q_2 P_2, \quad \Gamma \triangleq -P_2^{-1} P_{12}^T.$$

The existence of Q_2^{-1} and P_2^{-1} follows from the fact that (A_e, B_e, C_e) is minimal. See [1,2] for details. Note that $r = G^T \Gamma$. Clearly, Q, P, \hat{Q} , and \hat{P} are symmetric and nonnegative definite.

Next note that with the above definitions, (A.5) implies (4.2) and that (4.1) holds. Hence $r = G^T \Gamma$ is idempotent, i.e., $r^2 = r$. Sylvester's inequality yields (4.18). Note also that (4.6) and (4.7) hold.

The components of \mathcal{Q} and \mathcal{P} can be written in terms of $Q, P, \hat{Q}, \hat{P}, G$, and Γ as

$$Q_1 = Q + \hat{Q}, \quad P_1 = P + \hat{P}, \quad (\text{A.14})$$

$$Q_{12} = \hat{Q} \Gamma^T, \quad P_{12} = \hat{P} G^T, \quad (\text{A.15})$$

$$Q_2 = \Gamma \hat{Q} \Gamma^T, \quad P_2 = G \hat{P} G^T. \quad (\text{A.16})$$

Next note that by using (A.14)–(A.16), (A.7) becomes

$$C_e \hat{S} = L[I_n + \gamma^{-2}(Q + \hat{Q})P]G^T,$$

where

$$\hat{S} \triangleq I_{n_e} + \gamma^{-2} \Gamma \hat{Q} P G^T.$$

To prove that \hat{S} is invertible use (4.6) and (4.3) and note that

$$\begin{aligned} I_{n_e} + \gamma^{-2} \Gamma \hat{Q} P G^T &= I_{n_e} + \gamma^{-2} \Gamma \hat{Q} r^T P G^T \\ &= I_{n_e} + \gamma^{-2} (\Gamma \hat{Q} \Gamma^T) (G P G^T). \end{aligned}$$

Since $\Gamma \hat{Q} \Gamma^T$ and $G P G^T$ are nonnegative definite, their product has nonnegative eigenvalues. Thus each eigenvalue of $I_{n_e} + \gamma^{-2} \Gamma \hat{Q} P G^T$ is real and is greater than unity. Hence \hat{S} is invertible. Now note that by using (4.2) and (4.3) it can be shown that

$$G^T \hat{S}^{-1} \Gamma = S r.$$

The expressions (4.11), (4.12) and (4.13) follow from (A.6), (A.7), (4.8) and the definition of \mathcal{Q} by using the above identities. Next, computing either $\Gamma(\text{A.9}) - (\text{A.10})$ or $G(\text{A.12}) + (\text{A.13})$ yields (4.10). Substituting this expression for A_e into (A.8) – (A.13) it follows that (A.10) = $\Gamma(\text{A.9})$ and (A.13) = $G(\text{A.12})$. Thus, (A.10) and (A.13) are superfluous and can be omitted. Next, using (A.8) + $G^T \Gamma(\text{A.9})G - (\text{A.9})G - [(\text{A.9})G]^T$ and $G^T \Gamma(\text{A.9})G - (\text{A.9})G - [(\text{A.9})G]^T$ yields (4.14) and (4.16). Using (A.11) + $\Gamma^T G(\text{A.12})\Gamma - (\text{A.12})\Gamma - [(\text{A.12})\Gamma]^T$ and $\Gamma^T G(\text{A.12})\Gamma - (\text{A.12})\Gamma - [(\text{A.12})\Gamma]^T$ yields (4.15) and (4.17).

Finally, to prove the converse we use (4.10)–(4.18) to obtain (3.15) and (A.4) – (A.7). Let $A_e, B_e, C_e, G, \Gamma, \tau, Q, P, \hat{Q}, \hat{P}, \hat{Q}$ be as in the statement of Theorem 4.1 and define $Q_1, Q_{12}, Q_2, P_1, P_{12}, P_2$ by (A.14)–(A.16). Using (4.4), (4.11) and (4.12) it is easy to verify (A.6) and (A.7). Finally, substitute the definitions of $Q, P, \hat{Q}, \hat{P}, G, \Gamma$ and τ into (4.14)–(4.17) along with (4.2), (4.3), (4.6) and (4.7) to obtain (3.15) and (A.4). Finally, note that

$$\hat{Q} = \begin{bmatrix} Q & 0_{n \times n_e} \\ 0_{n_e \times n} & 0_{n_e} \end{bmatrix} + \begin{bmatrix} I_n \\ \Gamma \end{bmatrix} \hat{Q} \begin{bmatrix} I_n & \Gamma^T \end{bmatrix},$$

which shows that $\hat{Q} \geq 0$. \square

References

- [1] D.S. Bernstein and W.M. Haddad, "LQG Control with an H_∞ Performance Bound: A Riccati Equation Approach," *Proc. Amer. Contr. Conf.*, pp. 796–802, Atlanta, GA, June 1988.
- [2] D.S. Bernstein and D.C. Hyland, "The Optimal Projection Equations for Reduced-Order State Estimation," *IEEE Trans. Autom. Contr.*, Vol. AC-30, pp. 583–585, 1985.
- [3] R.W. Brockett, *Finite Dimensional Linear Systems*, Wiley, New York, 1970.
- [4] J.C. Doyle, K. Glover, P.P. Khargonekar, and B.A. Francis, "State-Space Solutions to Standard H_2 and H_∞ Control Problems," *Proc. Amer. Contr. Conf.*, pp. 1691–1696, Atlanta, GA, June 1988.
- [5] P.P. Khargonekar, I.R. Petersen and K. Zhou, "Robust Stabilization of Uncertain Systems and H^∞ Optimal Control," preprint.
- [6] I.R. Petersen, "Disturbance Attenuation and H^∞ Optimization: A Design Method Based on the Algebraic Riccati Equation," *IEEE Trans. Autom. Contr.*, Vol. AC-32, pp. 427–429, 1987.
- [7] C.R. Rao and S.K. Mitra, *Generalized Inverse of Matrices and Its Applications*, John Wiley and Sons, New York, 1971.
- [8] J.C. Willems, "Least Squares Stationary Optimal Control and the Algebraic Riccati Equation," *IEEE Trans. Autom. Contr.*, Vol. AC-16, pp. 621–634, 1971.
- [9] W.M. Wonham, *Linear Multivariable Control: A Geometric Approach*, Springer-Verlag, New York, 1979.

Submitted to IEEE TAC
November, 1987
Revised April, 1988

**LQG Control with an H_∞ Performance Bound:
A Riccati Equation Approach**

by

Dennis S. Bernstein
Harris Corporation
Government Aerospace Systems Division
MS 22/4848
Melbourne, FL 32902
(305) 729-2140

Wassim M. Haddad
Department of
Mechanical and Aerospace Engineering
Florida Institute of Technology
Melbourne, FL 32901

Abstract

An LQG control-design problem involving a constraint on H_∞ disturbance attenuation is considered. The H_∞ performance constraint is embedded within the optimization process by replacing the covariance Lyapunov equation by a Riccati equation whose solution leads to an upper bound on L_2 performance. In contrast to the pair of separated Riccati equations of standard LQG theory, the H_∞ -constrained gains are given by a coupled system of *three* modified Riccati equations. The coupling illustrates the breakdown of the separation principle for the H_∞ -constrained problem. Both full- and reduced-order design problems are considered with an H_∞ attenuation constraint involving both state and control variables. An algorithm is developed for the full-order design problem and illustrative numerical results are given.

Supported in part by the Air Force Office of Scientific Research under contracts F49620-86-C-0002 and F49620-87-C-0108.

1. Introduction

The fundamental differences between Wiener-Hopf-Kalman (WHK) control design (for example, LQG theory [1]) and H_∞ control theory [2-4] can be traced to the modeling and treatment of uncertain exogenous disturbances. As explained by Zames in the classic paper [2], LQG design is based upon a stochastic noise disturbance model possessing a fixed covariance (power spectral density), while H_∞ theory is predicated on a deterministic disturbance model consisting of bounded power (square-integrable) signals. Since LQG design utilizes a quadratic cost criterion, it follows from Plancherel's theorem that WHK theory strives to minimize the L_2 norm of the closed-loop frequency response, while H_∞ theory seeks to minimize the worst-case attenuation. For systems with poorly modeled disturbances which may possess significant power within arbitrarily small bandwidths, H_∞ is clearly appropriate, while, for systems with well-known disturbance power spectral densities, WHK design may be less conservative.

In addition to the fact that H_∞ design embodies many classical design objectives [5], it also presents a natural tool for modeling plant uncertainty in terms of normed H_∞ plant neighborhoods. In contrast, the H_2 topology has been shown in [6] to be too weak for a practical robustness theory, while the H_∞ norm is not only suitable for robust stabilization but is also conveniently submultiplicative. Within the WHK state space theory, however, the appropriate robustness model appears not to be a *nonparametric* normed plant neighborhood as in H_∞ theory, but rather a *parametric* uncertainty model. The principal technique for capturing the effects of real parameters within state space models is Lyapunov function theory (see, e.g., [7-16] and the references therein). Such structured uncertainties are difficult to capture nonconservatively within H_∞ theory except with specialized refinements ([17]).

In spite of the fundamental differences between WHK design and H_∞ theory, a significant connection was discovered in [18]. Specifically, Petersen observed that a modified algebraic Riccati equation developed for parameter-robust full-state-feedback control can also be utilized to yield controllers satisfying H_∞ disturbance attenuation bounds. This relationship was further explored in [19] where it was shown that Petersen's modified Riccati equation effectively yields the H_∞ -optimal full-state-feedback controller. This result is based upon the fact that quadratic stability (i.e., stability with a quadratic Lyapunov function) of the system

$$\dot{x} = (A + DFE)x, \quad \bar{\sigma}(F) \leq 1,$$

is equivalent to the stability of A and the small gain condition

$$\|E(sI - A)^{-1}D\|_{\infty} < 1.$$

The results of [19] thus solve the Standard Problem considered in [3,4] for the static full-state-feedback case.

The extension of these results to the Standard Problem for dynamic output-feedback compensation, however, was not given in [18,19]. Within the realm of quadratic robust stabilization, the dynamic output-feedback problem was addressed in [7]. The results of [7] involve a pair of decoupled modified Riccati equations along with an auxiliary inequality. Using different techniques, a more complete solution was obtained in [13,14] involving a *coupled* system of *three* modified Riccati equations for full-order dynamic compensation and a coupled system of *four* modified Riccati and Lyapunov equations in the fixed-order (i.e., reduced-order) case as in [20]. The results of [13,14] thus raise the following question: What is the relevance of this system of coupled design equations to the problem of H_{∞} disturbance attenuation via fixed-order compensation?

To begin to address this question the goal of the present paper is to *develop a design methodology for fixed-order L_2 optimal control which includes as a design constraint a bound on H_{∞} disturbance attenuation*. There are three principal motivations for developing such a theory. First, the results of [18,19] present full-state-feedback controllers whose form is directly analogous to the standard LQR solution. However, no L_2 interpretation was provided in [18,19] to explain this similarity. The present paper, however, provides an L_2 interpretation within the context of an H_{∞} design constraint. A novel feature of this mathematical formulation is the dual interpretation of the disturbances. That is, within the context of L_2 optimality the disturbances are interpreted as white noise signals while, simultaneously, for the purpose of H_{∞} attenuation the *very same disturbance signals have the alternative interpretation of deterministic L_2 functions*. This dual interpretation is unique to the present paper since stochastic modeling does not play a role in [18,19]. We also note recent results obtained in [21] which essentially show that the H_2 plant topology can be induced by imposing L_2 and L_{∞} topologies on the disturbance and output spaces, respectively. For further investigation into the relationships between L_2 and H_{∞} control, see [22].

The second motivation for our approach is the *simultaneous treatment of both L_2 and H_{∞} performance criteria which quantitatively demonstrates design tradeoffs*. Specifically, in order to enforce the H_{∞} constraint we derive an upper bound for the L_2 criterion. Minimization of this upper bound shows that the enforcement of an H_{∞} disturbance attenuation constraint leads directly

to an *increase* in the L_2 performance criterion. Although it would be preferable to minimize the L_2 criterion directly, this problem will be considered in a future paper.

The third motivation for our approach is to provide a characterization of *fixed-order* dynamic output-feedback controllers yielding specified disturbance attenuation. Existing optimal H_∞ design methods tend to yield relatively high-order controllers. Intuitively, solving the fixed-order design equations for progressively smaller H_∞ disturbance attenuation constraints should, in the limit, yield an H_∞ optimal controller over the class of fixed-order stabilizing controllers. Although our main result gives sufficient conditions, we also state hypotheses under which these conditions are also necessary (Proposition 4.1). It should also be noted that the inherent coupling among the modified Riccati equations shows that *the classical separation principle of LQG theory is not valid for the H_∞ -constrained full- and reduced-order design problems.*

Of course, since the present paper addresses the problem of fixed-order dynamic compensation, previous full-state feedback results ([18,19]) are not obtainable as a special case. However, applying the approach of the present paper to the full-state feedback problem yields results which are similar to those of [18,19].

Besides establishing connections with robust stabilizability in state space systems, an immediate benefit of the modified Riccati equation characterization of H_∞ -optimal controllers is the opportunity to develop novel computational algorithms for controller synthesis. To this end a continuation algorithm has been developed for solving the coupled system of three modified Riccati equations. In a numerical study (see Section 8) we have demonstrated convergence of the algorithm and reasonable computational efficiency. Homotopy methods were suggested for the coupled Riccati equations because of their demonstrated effectiveness in related problems which also involve coupled modified Riccati equations [23-25]. Since H_∞ control problems are solvable by established numerical methods [4], it should be stressed that the objective of these numerical studies is not to make direct comparisons with existing H_∞ synthesis algorithms, but only to demonstrate solvability of the coupled modified Riccati equations.

The contents of the paper are as follows. After presenting notation at the end of this section, the statement of the H_∞ -Constrained LQG Control Problem is given in Section 2. The principal result of this section (Lemma 2.1) shows that if the algebraic Lyapunov equation for the closed-loop covariance is replaced by a modified Riccati equation possessing a nonnegative-definite solution, then the closed-loop system is asymptotically stable, the H_∞ disturbance attenuation constraint is

satisfied, and the L_2 performance is bounded above by an "auxiliary" cost function. The problem of determining compensator gains which minimize this upper bound subject to the Riccati equation constraint is considered in Section 3 as the Auxiliary Minimization Problem. Necessary conditions for the Auxiliary Minimization Problem (Theorem 3.1) are given in the form of a coupled system of three modified Riccati equations. In Section 4 the necessary conditions of Theorem 3.1 are combined with Lemma 2.1 to yield sufficient conditions for closed-loop stability, H_∞ disturbance attenuation, and bounded L_2 performance. In Section 5 we specialize the results of Section 3 to the case in which the LQG weights are equal to the H_∞ weights. This serves to equalize the L_2 and H_∞ design aspects and, through a series of transformations, the results of Section 3 are recast in a simpler form. These results also clarify connections with [26]. To achieve further design flexibility, the reduced-order control-design problem is considered in Section 6. A simplified qualitative analysis of the full-order design equations is given in Section 7 to highlight important features with regard to existence and multiplicity of solutions. Finally, a numerical algorithm is presented in Section 8 along with illustrative numerical results. A series of designs are obtained to illustrate tradeoffs between the L_2 and H_∞ aspects and the conservatism of the L_2 performance bound. Although in the present paper the numerical results are limited to the case of full-order dynamic compensation, reduced-order designs have been obtained in [27] using Theorem 6.1.

Notation. Note: All matrices have real entries

| | |
|---|--|
| $\mathbb{R}, \mathbb{R}^{r \times s}, \mathbb{R}^r, \mathbb{E}$ | real numbers, $r \times s$ real matrices, $\mathbb{R}^{r \times 1}$, expected value |
| $I_r, (\cdot)^T, 0_{r \times s}, 0_r$ | $r \times r$ identity matrix, transpose, $r \times s$ zero matrix, $0_{r \times r}$ |
| $\text{tr}, \rho(\cdot)$ | trace, spectral radius |
| $\mathbb{S}^r, \mathbb{N}^r, \mathbb{P}^r$ | $r \times r$ symmetric, nonnegative-definite, positive-definite matrices |
| $Z_1 \leq Z_2, Z_1 < Z_2$ | $Z_2 - Z_1 \in \mathbb{N}^r, Z_2 - Z_1 \in \mathbb{P}^r, Z_1, Z_2 \in \mathbb{S}^r$ |
| $n, m, \ell, n_c, p, q, q_\infty; \tilde{n}$ | positive integers; $n + n_c$ ($n_c \leq n$) |
| x, u, y, x_c, \tilde{x} | $n, m, \ell, n_c, \tilde{n}$ -dimensional vectors |
| \tilde{x} | $\begin{bmatrix} x \\ x_c \end{bmatrix}$ |
| A, B, C | $n \times n, n \times m, \ell \times n$ matrices |
| A_c, B_c, C_c | $n_c \times n_c, n_c \times \ell, m \times n_c$ matrices |
| \tilde{A} | $\begin{bmatrix} A & BC_c \\ B_c C & A_c \end{bmatrix}$ |
| $w(\cdot)$ | p -dimensional standard white noise |
| D_1, D_2 | $n \times p, \ell \times p$ matrices; $D_1 D_2^T = 0$ |

| | |
|----------------------------|--|
| V_1, V_2 | $D_1 D_1^T, D_2 D_2^T; V_2 \in \mathbb{P}^l$ |
| \tilde{D} | $\begin{bmatrix} D_1 \\ B_c D_2 \end{bmatrix}$ |
| \tilde{V} | $\begin{bmatrix} V_1 & 0_{n \times n_c} \\ 0_{n_c \times n} & B_c V_2 B_c^T \end{bmatrix}$ |
| E_1, E_2 | $q \times n, q \times m$ matrices; $E_1^T E_2 = 0$ |
| \tilde{E} | $[E_1 \ E_2 C_c]$ |
| R_1, R_2 | $E_1^T E_1, E_2^T E_2; R_2 \in \mathbb{P}^m$ |
| \tilde{R} | $\begin{bmatrix} R_1 & 0_{n \times n_c} \\ 0_{n_c \times n} & C_c^T R_2 C_c \end{bmatrix} = \tilde{E}^T \tilde{E}$ |
| $E_{1\infty}, E_{2\infty}$ | $q_\infty \times n, q_\infty \times m$ matrices; $E_{1\infty}^T E_{2\infty} = 0$ |
| \tilde{E}_∞ | $[E_{1\infty} \ E_{2\infty} C_c]$ |
| $R_{1\infty}, R_{2\infty}$ | $E_{1\infty}^T E_{1\infty}, E_{2\infty}^T E_{2\infty}$ |
| \tilde{R}_∞ | $\begin{bmatrix} R_{1\infty} & 0_{n \times n_c} \\ 0_{n_c \times n} & C_c^T R_{2\infty} C_c \end{bmatrix} = \tilde{E}_\infty^T \tilde{E}_\infty$ |
| $\Sigma, \bar{\Sigma}$ | $B R_2^{-1} B^T, C^T V_2^{-1} C$ |
| β, γ | nonnegative constant, positive constant |

2. Statement of the Problem

In this section we introduce the LQG dynamic output-feedback control problem with constrained H_∞ disturbance attenuation between the plant and sensor disturbances and the state and control variables. Without the L_2 performance criterion the problem considered here essentially corresponds to the Standard Problem of [3,4]. For simplicity we restrict our attention to controllers of order $n_c = n$ only, i.e., controllers whose order is equal to the dimension of the plant. This constraint is removed in Section 6 where controllers of reduced order are considered. Hence, throughout Sections 2-5 the controller dimension n_c and closed-loop plant dimension $\tilde{n} \triangleq n + n_c$ should be interpreted as n and $2n$, respectively. Controllers of order greater than n are generally of less interest in practice and thus are not considered in this paper.

H_∞ -Constrained LQG Control Problem. Given the n th-order stabilizable and detectable plant

$$\dot{x}(t) = Ax(t) + Bu(t) + D_1 w(t), \quad (2.1)$$

$$y(t) = Cx(t) + D_2 w(t), \quad (2.2)$$

determine an n th-order dynamic compensator

$$\dot{x}_c(t) = A_c x_c(t) + B_c y(t), \quad (2.3)$$

$$u(t) = C_c x_c(t), \quad (2.4)$$

which satisfies the following design criteria:

- (i) the closed-loop system (2.1)–(2.4) is asymptotically stable, i.e., \tilde{A} is asymptotically stable;
- (ii) the $q_\infty \times p$ closed-loop transfer function

$$H(s) \triangleq \tilde{E}_\infty (sI_{\tilde{n}} - \tilde{A})^{-1} \tilde{D} \quad (2.5)$$

from $w(t)$ to $E_{1\infty}x(t) + E_{2\infty}u(t)$ satisfies the constraint

$$\|H(s)\|_\infty \leq \gamma, \quad (2.6)$$

where $\gamma > 0$ is a given constant; and

- (iii) the performance functional

$$J(A_c, B_c, C_c) \triangleq \lim_{t \rightarrow \infty} \mathbb{E}[x^T(t)R_1x(t) + u^T(t)R_2u(t)] \quad (2.7)$$

is minimized.

Note that the closed-loop system (2.1)–(2.4) can be written as

$$\dot{\tilde{x}}(t) = \tilde{A}\tilde{x}(t) + \tilde{D}w(t) \quad (2.8)$$

and that (2.7) becomes

$$J(A_c, B_c, C_c) = \lim_{t \rightarrow \infty} \mathbb{E}[(\tilde{E}\tilde{x}(t))^T(\tilde{E}\tilde{x}(t))] = \lim_{t \rightarrow \infty} \mathbb{E}[\tilde{x}^T(t)\tilde{R}\tilde{x}(t)]. \quad (2.9)$$

Remark 2.1. Since (A, B, C) is assumed to be stabilizable and detectable the set of n th-order stabilizing compensators is nonempty.

Remark 2.2. It is easy to show that the performance functional (2.7) is equivalent to the more familiar expression involving an averaged integral, i.e.,

$$J(A_c, B_c, C_c) = \lim_{t \rightarrow \infty} \frac{1}{t} \mathbb{E} \left\{ \int_0^t [x^T(s)R_1x(s) + u^T(s)R_2u(s)] ds \right\}.$$

Remark 2.3. For convenience we assume $D_1 D_2^T = 0$, which effectively implies that the plant disturbance and sensor noise are uncorrelated.

Remark 2.4. One may wish to consider a general L_2 optimization problem of the form $\min \|T - UQV\|_2$, where Q is a parameterization of stabilizing controllers. In this case, without a constraint on the MacMillan degree of Q , it may be possible to satisfy (2.6) with smaller values of γ .

Note that the problem statement involves both L_2 and H_∞ performance weights. In particular, the matrices R_1 and R_2 are the L_2 weights for the state and control variables. By introducing L_2 -weighted variables

$$z(t) = E_1 x(t), \quad v(t) = E_2 u(t),$$

the cost (2.7) can be written as

$$J(A_c, B_c, C_c) = \lim_{t \rightarrow \infty} \mathbb{E}[z^T(t)z(t) + v^T(t)v(t)].$$

For convenience we thus define $R_1 \triangleq E_1^T E_1$ and $R_2 \triangleq E_2^T E_2$ which appear in subsequent expressions. Although an L_2 cross-weighting term of the form $2x(t)^T R_{12} u(t)$ can also be included, we shall not do so here to facilitate the presentation.

For the H_∞ performance constraint, the transfer function (2.5) involves weighting matrices $E_{1\infty}$ and $E_{2\infty}$ for the state and control variables. The matrices $R_{1\infty} \triangleq E_{1\infty}^T E_{1\infty}$ and $R_{2\infty} \triangleq E_{2\infty}^T E_{2\infty}$ are thus the H_∞ counterparts of the L_2 weights R_1 and R_2 . Although we do not require that $R_{1\infty}$ and $R_{2\infty}$ be equal to R_1 and R_2 , we shall require in the next section that $R_{2\infty} = \beta^2 R_2$, where the nonnegative scalar β is a design variable. Finally, the condition $E_{1\infty}^T E_{2\infty} = 0$ precludes an H_∞ cross-weighting term which again facilitates the presentation.

Before continuing it is useful to note that if \tilde{A} is asymptotically stable for a given compensator (A_c, B_c, C_c) then the performance (2.7) is given by

$$J(A_c, B_c, C_c) = \text{tr } \tilde{Q} \tilde{R}, \quad (2.10)$$

where the steady-state closed-loop state covariance defined by

$$\tilde{Q} \triangleq \lim_{t \rightarrow \infty} \mathbb{E}[\tilde{x}(t)\tilde{x}^T(t)] \quad (2.11)$$

satisfies the $\tilde{n} \times \tilde{n}$ algebraic Lyapunov equation

$$0 = \tilde{A}\tilde{Q} + \tilde{Q}\tilde{A}^T + \tilde{V}. \quad (2.12)$$

Remark 2.5. Using (2.10) and (2.12) it can be shown that the L_2 cost criterion (2.7) can be written in terms of the L_2 norm of the impulse response of the closed-loop system. Specifically, writing \tilde{Q} satisfying (2.12) as

$$\tilde{Q} = \int_0^\infty e^{\tilde{A}t} \tilde{V} e^{\tilde{A}^T t} dt,$$

(2.10) becomes

$$J(A_c, B_c, C_c) = \int_0^\infty \|\tilde{E} e^{\tilde{A}t} \tilde{D}\|_F^2 dt,$$

where $\|\cdot\|_F$ denotes the Frobenius matrix norm.

The key step in enforcing the disturbance attenuation constraint (2.6) is to replace the algebraic Lyapunov equation (2.12) by an algebraic Riccati equation which overbounds the closed-loop steady-state covariance. Justification for this technique is provided by the following result.

Lemma 2.1. Let (A_c, B_c, C_c) be given and assume there exists $Q \in \mathbb{R}^{\tilde{n} \times \tilde{n}}$ satisfying

$$Q \in \mathbb{N}^{\tilde{n}} \quad (2.13)$$

and

$$0 = \tilde{A}Q + Q\tilde{A}^T + \gamma^{-2}Q\tilde{R}_\infty Q + \tilde{V}. \quad (2.14)$$

Then

$$(\tilde{A}, [\gamma^{-2}Q\tilde{R}_\infty Q + \tilde{V}]^{\frac{1}{2}}) \text{ is stabilizable} \quad (2.15)$$

if and only if

$$\tilde{A} \text{ is asymptotically stable.} \quad (2.16)$$

In this case,

$$\|H(s)\|_\infty \leq \gamma \quad (2.17)$$

and

$$\tilde{Q} \leq Q. \quad (2.18)$$

Consequently,

$$J(A_c, B_c, C_c) \leq J(A_c, B_c, C_c, Q), \quad (2.19)$$

where

$$J(A_c, B_c, C_c, Q) \triangleq \text{tr } Q\tilde{R}. \quad (2.20)$$

Proof. Using the assumed existence of a nonnegative-definite solution to (2.14) and the stabilizability condition (2.15), it follows from the dual of Lemma 12.2 of [28] that \tilde{A} is asymptotically

stable. Conversely, if \tilde{A} is asymptotically stable then (2.15) holds. The proof of (2.17) follows from a standard manipulation of (2.14); for details see Lemma 1 of [29]. To prove (2.18) subtract (2.12) from (2.14) to obtain

$$0 = \tilde{A}(\mathcal{Q} - \tilde{\mathcal{Q}}) + (\mathcal{Q} - \tilde{\mathcal{Q}})\tilde{A}^T + \gamma^{-2}\mathcal{Q}\tilde{R}_\infty\mathcal{Q}, \quad (2.21)$$

which, since \tilde{A} is asymptotically stable, is equivalent to

$$\mathcal{Q} - \tilde{\mathcal{Q}} = \int_0^\infty e^{\tilde{A}t}[\gamma^{-2}\mathcal{Q}\tilde{R}_\infty\mathcal{Q}]e^{\tilde{A}^T t} dt \geq 0. \quad (2.22)$$

Finally, (2.19) follows immediately from (2.18). \square

Remark 2.6. Note that (2.15) is actually a closed-loop "disturbability" condition which is not concerned with control as such. This condition guarantees that the system does not possess unstable undisturbed modes. Of course, if \tilde{V} is positive definite or (\tilde{A}, \tilde{D}) is controllable, then (2.15) is satisfied.

Lemma 2.1 shows that the H_∞ disturbance attenuation constraint is automatically enforced when a nonnegative-definite solution to (2.14) is known to exist. Furthermore, the solution \mathcal{Q} provides an upper bound for the *actual* closed-loop state covariance $\tilde{\mathcal{Q}}$ along with a bound on the L_2 performance criterion. Next, we present a partial converse of Lemma 2.1 which guarantees the existence of a nonnegative-definite solution to (2.14) when (2.17) is satisfied.

Lemma 2.2. Let (A_c, B_c, C_c) be given, suppose \tilde{A} is asymptotically stable, and assume the disturbance attenuation constraint (2.17) is satisfied. Then there exists a unique nonnegative-definite solution \mathcal{Q} satisfying (2.14). Furthermore, $(\tilde{A} + \gamma^{-2}\mathcal{Q}\tilde{R}_\infty, \tilde{V}^{\frac{1}{2}})$ is stabilizable if and only if $\tilde{A} + \gamma^{-2}\mathcal{Q}\tilde{R}_\infty$ is asymptotically stable.

Proof. The result is an immediate consequence of [30], using Theorems 3 and 2, pp. 150 and 167, along with the dual version of Lemma 12.2 of [28]. \square

Remark 2.7. To further clarify the relationships between the L_2 and H_∞ aspects of the problem, we note that the closed-loop system can be represented by two possibly different transfer functions. Specifically, with respect to the L_2 cost criterion, the closed-loop transfer function between disturbances and controlled variables is given by the triple $(\tilde{A}, \tilde{D}, \tilde{E})$ while for the H_∞ constraint the closed-loop transfer function (2.5) corresponds to the triple $(\tilde{A}, \tilde{D}, \tilde{E}_\infty)$.

Finally, it can be shown that the closed-loop Riccati equation (2.14) also enforces a constraint on the norm of the Hankel operator corresponding to the closed-loop system $(\tilde{A}, \tilde{D}, \tilde{E}_\infty)$ when \mathcal{Q}

is positive definite. Thus, let $\tilde{P} \in \mathbb{R}^n$ denote the solution to

$$0 = \tilde{A}^T \tilde{P} + \tilde{P} \tilde{A} + \tilde{R}_\infty \quad (2.23)$$

and note that \tilde{P} and \tilde{Q} (satisfying (2.12)) are the observability and controllability Gramians, respectively, of the system $(\tilde{A}, \tilde{D}, \tilde{E}_\infty)$. As shown in [31], the norm of the Hankel operator corresponding to $(\tilde{A}, \tilde{D}, \tilde{E}_\infty)$ is given by $\lambda_{\max}^{\frac{1}{2}}(\tilde{P}\tilde{Q})$.

Proposition 2.1. Suppose there exists $Q \in \mathbb{R}^n$ satisfying (2.14) and such that (2.15) or, equivalently, (2.16) holds. Then

$$\lambda_{\max}^{\frac{1}{2}}(\tilde{P}\tilde{Q}) \leq \gamma. \quad (2.24)$$

Proof. Since Q is assumed to be invertible, (2.14) is equivalent to

$$0 = \gamma^2 \tilde{A}^T Q^{-1} + \gamma^2 Q^{-1} \tilde{A} + \gamma^2 Q^{-1} \tilde{V} Q^{-1} + \tilde{R}_\infty. \quad (2.25)$$

Subtracting (2.23) from (2.25) shows that $\gamma^2 Q^{-1} - \tilde{P} \geq 0$, or, equivalently, $\gamma^2 I_n \geq Q^{\frac{1}{2}} \tilde{P} Q^{\frac{1}{2}}$. Thus,

$$\gamma^2 \geq \lambda_{\max}(Q^{\frac{1}{2}} \tilde{P} Q^{\frac{1}{2}}) = \lambda_{\max}(\tilde{P}^{\frac{1}{2}} Q \tilde{P}^{\frac{1}{2}}) \geq \lambda_{\max}(\tilde{P}^{\frac{1}{2}} \tilde{Q} \tilde{P}^{\frac{1}{2}}) = \lambda_{\max}(\tilde{P}\tilde{Q}),$$

which yields (2.24). \square

3. The Auxiliary Minimization Problem and Necessary Conditions for Optimality

As discussed in the previous section, the replacement of (2.12) by (2.14) enforces the H_∞ disturbance attenuation constraint and yields an upper bound for the L_2 performance criterion. That is, given a compensator (A_c, B_c, C_c) for which there exists a nonnegative-definite solution to (2.14), the *actual* L_2 performance $J(A_c, B_c, C_c)$ of the compensator is guaranteed to be no worse than the bound given by $J(A_c, B_c, C_c, Q)$. Hence, $J(A_c, B_c, C_c, Q)$ can be interpreted as an *auxiliary cost* which leads to the following mathematical programming problem.

Auxiliary Minimization Problem. Determine (A_c, B_c, C_c, Q) which minimizes $J(A_c, B_c, C_c, Q)$ subject to (2.13) and (2.14).

It follows from Lemma 2.1 that the satisfaction of (2.13) and (2.14) along with the generic condition (2.15) leads to 1) closed-loop stability; 2) prespecified H_∞ performance attenuation; and 3) an upper bound for the L_2 performance criterion. Hence it remains to determine (A_c, B_c, C_c) which minimizes $J(A_c, B_c, C_c, Q)$ and thus provides an optimized bound for the actual L_2 performance

$J(A_e, B_e, C_e)$. Rigorous derivation of the necessary conditions for the Auxiliary Minimization Problem requires additional technical assumptions. Specifically, we restrict (A_e, B_e, C_e, Q) to the open set

$$\begin{aligned} \mathcal{X} \triangleq \{ & (A_e, B_e, C_e, Q) : Q \in \mathbb{P}^n, \tilde{A} + \gamma^{-2} Q \tilde{R}_\infty \text{ is asymptotically stable,} \\ & \text{and } (A_e, B_e, C_e) \text{ is controllable and observable} \}. \end{aligned} \quad (3.1)$$

Remark 3.1. The set \mathcal{X} constitutes sufficient conditions under which the Lagrange multiplier technique is applicable to the Auxiliary Minimization Problem. Specifically, the requirement that Q be positive definite replaces (2.13) by an open set constraint, the stability of $\tilde{A} + \gamma^{-2} Q \tilde{R}_\infty$ serves as a normality condition, and (A_e, B_e, C_e) minimal is a nondegeneracy condition. Note that the stabilizability condition (2.15) and stability condition (2.16) play no role in determining solutions of the Auxiliary Minimization Problem.

The following result presents the necessary conditions for optimality in the Auxiliary Minimization Problem. The proof of this result is given in Appendix A as a special case of the corresponding result for reduced-order dynamic compensation considered in Section 6. As mentioned previously, we assume that $R_{2\infty} = \beta^2 R_2$. Furthermore, for arbitrary $\hat{Q}, P \in \mathbb{N}^n$ define

$$S \triangleq (I_n + \beta^2 \gamma^{-2} \hat{Q} P)^{-1}. \quad (3.2)$$

Since the eigenvalues of $\hat{Q} P$ coincide with the eigenvalues of the nonnegative-definite matrix $P^{\frac{1}{2}} \hat{Q} P^{\frac{1}{2}}$, it follows that $\hat{Q} P$ has nonnegative eigenvalues. Thus, the eigenvalues of $I_n + \beta^2 \gamma^{-2} \hat{Q} P$ are all greater than one so that the above inverse exists.

Theorem 3.1. If $(A_e, B_e, C_e, Q) \in \mathcal{X}$ solves the Auxiliary Minimization Problem then there exist $Q, P, \hat{Q} \in \mathbb{N}^n$ such that

$$A_e = A - Q \tilde{E} - \Sigma P S + \gamma^{-2} Q R_{1\infty}, \quad (3.3)$$

$$B_e = Q C^T V_2^{-1}, \quad (3.4)$$

$$C_e = -R_2^{-1} B^T P S, \quad (3.5)$$

$$Q = \begin{bmatrix} Q + \hat{Q} & \hat{Q} \\ \hat{Q} & \hat{Q} \end{bmatrix}, \quad (3.6)$$

and such that Q, P, \hat{Q} satisfy

$$0 = A Q + Q A^T + V_1 + \gamma^{-2} Q R_{1\infty} Q - Q \tilde{E} Q, \quad (3.7)$$

$$0 = (A + \gamma^{-2}[Q + \hat{Q}]R_{1\infty})^T P + P(A + \gamma^{-2}[Q + \hat{Q}]R_{1\infty}) + R_1 - S^T P \Sigma P S, \quad (3.8)$$

$$0 = (A - \Sigma P S + \gamma^{-2}Q R_{1\infty})\hat{Q} + \hat{Q}(A - \Sigma P S + \gamma^{-2}Q R_{1\infty})^T + \gamma^{-2}\hat{Q}(R_{1\infty} + \beta^2 S^T P \Sigma P S)\hat{Q} + Q \bar{\Sigma} Q. \quad (3.9)$$

Furthermore, the auxiliary cost is given by

$$J(A_c, B_c, C_c, Q) = \text{tr}[(Q + \hat{Q})R_1 + \hat{Q}S^T P \Sigma P S]. \quad (3.10)$$

Conversely, if there exist $Q, P, \hat{Q} \in \mathbb{N}^n$ satisfying (3.7)–(3.9), then (A_c, B_c, C_c, Q) given by (3.3)–(3.6) satisfies (2.13) and (2.14) with auxiliary cost (2.20) given by (3.10).

Remark 3.2. If Q and \hat{Q} are nonnegative definite then the fact that (2.13) is satisfied can easily be seen by writing Q as

$$Q = \begin{bmatrix} Q & 0_n \\ 0_n & 0_n \end{bmatrix} + \begin{bmatrix} \hat{Q}^{\frac{1}{2}} \\ \hat{Q}^{\frac{1}{2}} \end{bmatrix} \begin{bmatrix} \hat{Q}^{\frac{1}{2}} \\ \hat{Q}^{\frac{1}{2}} \end{bmatrix}^T.$$

Remark 3.3. Setting $\beta = 0$ or, equivalently, $E_{2\infty} = 0$, specializes Theorem 3.1 to the “cheap” H_∞ control case in which H_∞ attenuation between disturbances and controls is not constrained. In this case $S = I_n$, Q is given by (3.6), and

$$A_c = A - Q \bar{\Sigma} - \Sigma P + \gamma^{-2}Q R_{1\infty}, \quad (3.11)$$

$$B_c = Q C^T V_2^{-1}, \quad (3.12)$$

$$C_c = -R_2^{-1} B^T P, \quad (3.13)$$

where Q satisfies (3.7), and equations (3.8) and (3.9) become

$$0 = (A + \gamma^{-2}[Q + \hat{Q}]R_{1\infty})^T P + P(A + \gamma^{-2}[Q + \hat{Q}]R_{1\infty}) + R_1 - P \Sigma P, \quad (3.14)$$

$$0 = (A - \Sigma P + \gamma^{-2}Q R_{1\infty})\hat{Q} + \hat{Q}(A - \Sigma P + \gamma^{-2}Q R_{1\infty})^T + \gamma^{-2}\hat{Q}R_{1\infty}\hat{Q} + Q \bar{\Sigma} Q. \quad (3.15)$$

Finally, the auxiliary cost reduces to

$$J(A_c, B_c, C_c, Q) = \text{tr}[(Q + \hat{Q})R_1 + \hat{Q}P \Sigma P]. \quad (3.16)$$

Numerical solution of equations (3.7), (3.14) and (3.15) is discussed in Section 8.

Theorem 3.1 presents necessary conditions for the Auxiliary Minimization Problem which explicitly synthesize extremal controllers (A_c, B_c, C_c) . These necessary conditions comprise a system

of three modified algebraic Riccati equations in variables Q, P and \hat{Q} . The Q and P equations are similar to the estimator and regulator Riccati equations of LQG theory, while the \hat{Q} equation has no counterpart in the standard theory. Note that the Q equation is decoupled from the P and \hat{Q} equations and thus can be solved independently. The P equation, however, depends on Q . Thus, regulator/estimator separation only holds in one direction which clearly shows that *the certainty equivalence principle is no longer valid for the design problem under consideration*. Furthermore, since the P and \hat{Q} equations are coupled, they must be solved simultaneously. Finally, note that if the H_∞ disturbance attenuation constraint is sufficiently relaxed, i.e., $\gamma \rightarrow \infty$, then the P equation becomes decoupled from the \hat{Q} equation and thus the \hat{Q} equation becomes superfluous. Furthermore, the remaining Q and P equations separate and coincide with the standard LQG result.

4. Sufficient Conditions for H_∞ Disturbance Attenuation

In this section we combine Lemma 2.1 with the converse of Theorem 3.1 to obtain our main result guaranteeing closed-loop stability, H_∞ disturbance attenuation, and an optimized L_2 performance bound.

Theorem 4.1. Suppose there exist $Q, P, \hat{Q} \in \mathbb{N}^n$ satisfying (3.7)–(3.9). Then, with (A_e, B_e, C_e, Q) given by (3.3)–(3.6), $(\tilde{A}, [\gamma^{-2} Q \tilde{R}_\infty Q + \tilde{V}]^{\frac{1}{2}})$ is stabilizable if and only if \tilde{A} is asymptotically stable. In this case, the closed-loop transfer function $H(s)$ satisfies the H_∞ attenuation constraint

$$\|H(s)\|_\infty \leq \gamma, \quad (4.1)$$

and the L_2 performance criterion (2.7) satisfies the bound

$$J(A_e, B_e, C_e) \leq \text{tr}[(Q + \hat{Q})R_1 + \hat{Q}S^T P \Sigma P S]. \quad (4.2)$$

Proof. The converse portion of Theorem 3.1 implies that Q given by (3.6) satisfies (2.13) and (2.14) with auxiliary cost given by (3.10). It now follows from Lemma 2.1 that the stabilizability condition (2.15) is equivalent to the asymptotic stability of \tilde{A} , the H_∞ disturbance attenuation constraint (2.17) holds, and the performance bound (2.19), which is equivalent to (4.2), holds. \square

Remark 4.1. In applying Theorem 4.1 it is not actually necessary to check (2.15) which holds generically. Rather, it suffices to check the stability of \tilde{A} directly which is guaranteed to be equivalent to (2.15).

In applying Theorem 4.1 the principal issue concerns conditions on the problem data under which the coupled Riccati equations (3.7)–(3.9) possess nonnegative-definite solutions. Clearly, for γ sufficiently large, (3.7)–(3.9) approximate the standard LQG result so that existence and uniqueness are assured. The important case of interest, however, involves small γ so that significant H_∞ disturbance attenuation is enforced. Thus, if (4.1) can be satisfied for a given $\gamma > 0$, it is of interest to know whether one such controller can be obtained by solving (3.7)–(3.9). Lemma 2.2 guarantees that (2.14) possesses a solution for any controller satisfying (4.1). Thus our sufficient condition will also be *necessary* so long as the Auxiliary Minimization Problem possesses at least one extremal over \mathcal{X} . When this is the case we have the following immediate result.

Proposition 4.1. Let γ^* denote the infimum of $\|H(s)\|_\infty$ over all stabilizing n th-order dynamic compensators and suppose that the Auxiliary Minimization Problem has a solution for all $\gamma > \gamma^*$. Then for all $\gamma > \gamma^*$ there exist $Q, P, \tilde{Q} \in \mathbb{R}^n$ satisfying (3.7)–(3.9).

Unlike the standard LQG result involving a pair of separated Riccati equations, the new result enforcing H_∞ disturbance attenuation involves a nonstandard coupled system of *three* modified Riccati equations. The asymmetry of these equations suggests the possibility of a *dual* result in which the modifications to the standard P and Q Riccati equations are reversed. Such a dual result will generally be different from Theorem 4.1 and thus will yield an improved bound for particular problems. This point was demonstrated in [16] for the problem of robust performance analysis. Due to space limitations, however, we give only a brief outline of the dual H_∞ results. Note that $J(A_c, B_c, C_c)$ given by (2.10) is also given by

$$J(A_c, B_c, C_c) = \text{tr } \tilde{P}\tilde{V}, \quad (4.3)$$

where $\tilde{P} \in \mathbb{R}^n$ is the unique solution to (2.23). Next, utilizing (4.3) in place of (2.10), the H_∞ disturbance attenuation constraint (2.6) can now be enforced by replacing (2.23) by the Riccati equation

$$0 = \tilde{A}^T \tilde{P} + \tilde{P} \tilde{A} + \gamma^{-2} \tilde{P} \tilde{V} \tilde{P} + \tilde{R}_\infty. \quad (4.4)$$

Note that (4.4) is merely the dual of (2.14). We also require the condition dual to (2.15) given by

$$([\gamma^{-2} \tilde{P} \tilde{V} \tilde{P} + \tilde{R}_\infty]^{\frac{1}{2}}, \tilde{A}) \text{ is detectable.} \quad (4.5)$$

Once again, the sufficient conditions for H_∞ disturbance attenuation involve a coupled system of three modified Riccati equations dual to (3.7)–(3.9). Similar remarks apply to the reduced-order case given by Theorem 6.1 below.

5. Alternative Forms of the Riccati Equations

The purpose of this section is to draw connections with recent results obtained in [26]. As shown in Theorem 4.1, the Riccati equations (3.7)–(3.9) provide sufficient conditions for explicitly synthesizing controllers (A_c, B_c, C_c) satisfying an H_∞ performance bound. In this section we specialize Theorem 3.1 by setting the LQG weights equal to the H_∞ weights, i.e., $R_1 = R_{1\infty}$ and $\beta = 1$. This specialization leads to considerable simplification by equalizing the L_2 and H_∞ design aspects. In this case it turns out that the Riccati equations (3.7)–(3.9) can be transformed to simpler forms which are similar to the results given in [26]. To state the results we require some additional notation and a lemma concerning transformations of a pair of nonnegative-definite matrices.

Lemma 5.1. Let $P, Q \in \mathbb{R}^n$. Then the following statements hold.

- (i) PQ has nonnegative eigenvalues.
- (ii) $\rho(PQ) \leq (<) \gamma^{-2}$ if and only if $P^{\frac{1}{2}} Q P^{\frac{1}{2}} \leq (<) \gamma^{-2} I_n$. Furthermore, if P is positive definite then $\rho(PQ) \leq (<) \gamma^{-2}$ if and only if $Q \leq (<) \gamma^{-2} P^{-1}$.
- (iii) Define $Z \triangleq P(I_n + \gamma^{-2} PQ)^{-1}$. Then Z is nonnegative definite,

$$P(I_n + \gamma^{-2} QP)^{-1} = (I_n + \gamma^{-2} PQ)^{-1} P, \quad (5.1)$$

and $\rho(ZQ) < \gamma^{-2}$ or, equivalently, $Q^{\frac{1}{2}} Z Q^{\frac{1}{2}} < \gamma^{-2} I_n$. Furthermore, Z is positive definite if and only if P is positive definite. In this case,

$$Z = (P^{-1} + \gamma^{-2} Q)^{-1}. \quad (5.2)$$

- (iv) Suppose $\rho(PQ) < \gamma^{-2}$ and define $Y \triangleq P(I_n - \gamma^{-2} QP)^{-1}$. Then Y is nonnegative definite,

$$P(I_n - \gamma^{-2} QP)^{-1} = (I_n - \gamma^{-2} PQ)^{-1} P, \quad (5.3)$$

and

$$QP(I_n - \gamma^{-2} QP)^{-1} = (I_n - \gamma^{-2} QP)^{-1} QP. \quad (5.4)$$

Furthermore, P is positive definite if and only if Y is positive definite. In this case,

$$Y = (P^{-1} - \gamma^{-2} Q)^{-1}. \quad (5.5)$$

Proof. The results are easily proven and thus the details are omitted. \square

Proposition 5.1. Let $R_1 = R_{1\infty}$, $\beta = 1$, suppose there exist $Q, \hat{Q} \in \mathbb{N}^n$ and $Z \in \mathbb{P}^n$ satisfying

$$0 = AQ + QA^T + V_1 + \gamma^{-2}QR_1Q - Q\bar{E}Q, \quad (5.6)$$

$$0 = (A + \gamma^{-2}QR_1)^T Z + Z(A + \gamma^{-2}QR_1) + R_1 - Z\Sigma Z + \gamma^{-2}ZQ\bar{E}QZ, \quad (5.7)$$

$$0 = (A - \Sigma Z + \gamma^{-2}QR_1)\hat{Q} + \hat{Q}(A - \Sigma Z + \gamma^{-2}QR_1)^T + \gamma^{-2}\hat{Q}(R_1 + Z\Sigma Z)\hat{Q} + Q\bar{E}Q, \quad (5.8)$$

$$\rho(Z\hat{Q}) < \gamma^{-2}, \quad (5.9)$$

and let (A_c, B_c, C_c, Q) be given by

$$A_c = A - Q\bar{E} - \Sigma Z + \gamma^{-2}QR_1, \quad (5.10)$$

$$B_c = QC^TV_2^{-1}, \quad (5.11)$$

$$C_c = -R_2^{-1}B^TZ, \quad (5.12)$$

$$Q = \begin{bmatrix} Q + \hat{Q} & \hat{Q} \\ \hat{Q} & \hat{Q} \end{bmatrix}. \quad (5.13)$$

Then $(\tilde{A}, [\gamma^{-2}Q\tilde{R}Q + \tilde{V}]^{\frac{1}{2}})$ is stabilizable if and only if \tilde{A} is asymptotically stable. In this case, the closed-loop transfer function $H(s)$ satisfies the H_∞ disturbance attenuation constraint

$$\|H(s)\|_\infty \leq \gamma, \quad (5.14)$$

and the L_2 performance criterion (2.7) satisfies the bound

$$J(A_c, B_c, C_c) \leq \text{tr}[(Q + \hat{Q})R_1 + \hat{Q}Z\Sigma Z]. \quad (5.15)$$

Proof. We need only show that (5.6)–(5.9) imply (3.7)–(3.9). Clearly, (3.7) is a restatement of (5.6). Since Z is positive definite and $\rho(Z\hat{Q}) < \gamma^{-2}$, it is possible by Lemma 5.1 (iv) to define $P \triangleq Z(I_n - \gamma^{-2}\hat{Q}Z)^{-1} = (Z^{-1} - \gamma^{-2}\hat{Q})^{-1}$. Furthermore, P is positive definite and it follows that $Z = PS$. Thus (5.8) implies (3.9). To obtain (3.8), form $P[Z^{-1}(5.7)Z^{-1} - \gamma^{-2}(5.8)]P$ and combine terms. \square

Remark 5.1. It is easy to see that the proof of Proposition 5.1 can be reversed. That is, when P satisfying (3.8) is positive definite define $Z \triangleq (P^{-1} + \gamma^{-2}\hat{Q})^{-1}$ which leads to (5.7) and (5.8) using Lemma 5.1 (iii). Thus, when P is positive definite, the form of the equations (5.6)–(5.9) represents no loss of generality.

Proposition 5.2. Let $R_1 = R_{1\infty}, \beta = 1$, suppose there exist $Q, \hat{Q} \in \mathbb{N}^n$ and $P_0 \in \mathbb{P}^n$ satisfying

$$0 = AQ + QA^T + V_1 + \gamma^{-2}QR_1Q - Q\bar{\Sigma}Q, \quad (5.16)$$

$$0 = A^TP_0 + P_0A + R_1 + \gamma^{-2}P_0V_1P_0 - P_0\Sigma P_0, \quad (5.17)$$

$$0 = [A - \Sigma P_0(I_n - \gamma^{-2}QP_0)^{-1} + \gamma^{-2}QR_1]\hat{Q} + \hat{Q}[A - \Sigma P_0(I_n - \gamma^{-2}QP_0)^{-1} + \gamma^{-2}QR_1]^T \\ + \gamma^{-2}\hat{Q}[R_1 + P_0(I_n - \gamma^{-2}QP_0)^{-1}\Sigma(I_n - \gamma^{-2}P_0Q)^{-1}P_0]\hat{Q} + Q\bar{\Sigma}Q, \quad (5.18)$$

$$\rho[(Q + \hat{Q})P_0] < \gamma^{-2}, \quad (5.19)$$

and let (A_c, B_c, C_c, Q) be given by

$$A_c = A - Q\bar{\Sigma} - \Sigma P_0(I_n - \gamma^{-2}QP_0)^{-1} + \gamma^{-2}QR_1, \quad (5.20)$$

$$B_c = QC^TV_2^{-1}, \quad (5.21)$$

$$C_c = -R_2^{-1}B^TP_0(I_n - \gamma^{-2}QP_0)^{-1}, \quad (5.22)$$

$$Q = \begin{bmatrix} Q + \hat{Q} & \hat{Q} \\ \hat{Q} & \hat{Q} \end{bmatrix}. \quad (5.23)$$

Then $(\tilde{A}, [\gamma^{-2}Q\tilde{R}Q + \tilde{V}]^{\frac{1}{2}})$ is stabilizable if and only if \tilde{A} is asymptotically stable. In this case, the closed-loop transfer function $H(s)$ satisfies the H_∞ disturbance attenuation constraint

$$\|H(s)\|_\infty \leq \gamma, \quad (5.24)$$

and the L_2 performance criterion (2.7) satisfies the bound

$$J(A_c, B_c, C_c) \leq \text{tr}[(Q + \hat{Q})R_1 + \hat{Q}P_0(I_n - \gamma^{-2}QP_0)^{-1}\Sigma(I_n - \gamma^{-2}P_0Q)^{-1}P_0]. \quad (5.25)$$

Proof. As in the proof of Proposition 5.1, it need only be shown that (5.16)–(5.19) imply (3.7)–(3.9). Hence define $P = [P_0^{-1} - \gamma^{-2}(Q + \hat{Q})]^{-1}$ and form the equation $P[P_0^{-1}(5.17)P_0^{-1} - \gamma^{-2}(5.16) - \gamma^{-2}(5.18)]P$ to obtain (3.8). Equations (3.7) and (3.9) are immediate. \square

Remark 5.2. To clarify the relationships among (3.7)–(3.9), (5.6)–(5.8), and (5.16)–(5.18), we tabulate the transformations involving P, Z and P_0 :

$$P = (Z^{-1} - \gamma^{-2}\hat{Q})^{-1} = [P_0^{-1} - \gamma^{-2}(Q + \hat{Q})]^{-1}, \quad (5.26)$$

$$Z = (P^{-1} + \gamma^{-2}\hat{Q})^{-1} = (P_0^{-1} - \gamma^{-2}Q)^{-1}, \quad (5.27)$$

$$P_0 = (Z^{-1} + \gamma^{-2}Q)^{-1} = [P^{-1} + \gamma^{-2}(Q + \hat{Q})]^{-1}. \quad (5.28)$$

Remark 5.3. It is important to note that numerically solving equations (5.6)–(5.8) and (5.16)–(5.18) does not require that Z and P_0 be invertible. As shown in the Proofs of Propositions 5.1 and 5.2 the positive definiteness assumptions are used to construct equation (3.8).

Remark 5.4. Note that the gains (5.10)–(5.12) and (5.16)–(5.18) of Propositions 5.1 and 5.2 are *independent* of the matrix \hat{Q} satisfying equations (5.8) and (5.18). Nevertheless, equations (5.8) and (5.18) *must* have a solution in order to enforce the solvability of (2.14) which implies that the H_∞ constraint is satisfied. Thus our result does not yield a guarantee of H_∞ performance unless (5.8) and (5.18) can be solved numerically. Of course, \hat{Q} is also required to evaluate the L_2 performance bound (5.15) or (5.25). Note that the solutions of Q and P_0 of (5.16) and (5.17) are analogous to the matrices X_∞ and Y_∞ of [26]. Finally, note that (5.19) implies that $\rho(QP_0) < \gamma^{-2}$ which is essentially condition 5.2 (iii) of [26].

Remark 5.5. The transformations (5.26)–(5.28) used to obtain the form of the equations given by Propositions 5.1 and 5.2 depend strongly upon the assumptions $R_{1\infty} = R_1$ and $\beta = 1$. That is, if either $R_{1\infty} \neq R_1$ or $\beta \neq 1$ then these transformations cannot be carried out. Thus, although (5.6)–(5.8) and (5.16)–(5.18) possess numerical advantages over (3.7)–(3.9), these alternative forms exist only in the very special case in which the L_2 and H_∞ weights are equalized. Moreover, in the presence of parameter uncertainties ([13,15]) or nonstrictly proper controller design ([32]), such transformations seem to be precluded.

Remark 5.6. It is interesting to note that equations (5.6) and (5.7) with controller gains (5.10)–(5.12) are already known since they are identical to the optimality conditions of the exponential-quadratic-Gaussian problem treated in [33]. Specifically, see equations (3.1) and (4.1) on pages 603 and 609, respectively. As shown in [33], minimizing the criterion

$$J = \lim_{t \rightarrow \infty} \mathbb{E} \mu e^{\frac{\mu}{2} (x^T R_1 x + u^T R_2 u)}$$

leads to the pair of modified Riccati equations (5.6) and (5.7) with γ^{-2} replaced by μ . This implies that with equation (5.8) the exponential-of-quadratic design problem effectively enforces a bound of $\mu^{-\frac{1}{2}}$ on the H_∞ norm of the closed-loop transfer function. What remains to be achieved then is a deeper understanding of this connection. For related references, see [34,35].

6. Extensions to Reduced-Order Dynamic Compensation

In this section we extend Theorem 4.1 by expanding the formulation of Section 3 to allow the compensator to be of fixed dimension n_c which may be less than the plant order n . Hence, in this section define $\tilde{n} = n + n_c$, where $n_c \leq n$. As in [20] this constraint leads to an oblique projection which introduces additional coupling in the design equations along with an additional equation. The following lemma is required.

Lemma 6.1. Let $\hat{Q}, \hat{P} \in \mathbb{R}^n$ and suppose $\text{rank } \hat{Q}\hat{P} = n_c$. Then there exist $n_c \times n$ G, Γ and $n_c \times n_c$ invertible M , unique except for a change of basis in \mathbb{R}^{n_c} , such that

$$\hat{Q}\hat{P} = G^T M \Gamma, \quad \Gamma G^T = I_{n_c}. \quad (6.1), (6.2)$$

Furthermore, the $n \times n$ matrices

$$\tau \triangleq G^T \Gamma, \quad \tau_{\perp} \triangleq I_n - \tau \quad (6.3), (6.4)$$

are idempotent and have rank n_c and $n - n_c$, respectively.

Proof. Conditions (6.1)–(6.4) are a direct consequence of Theorem 6.2.5 of [36]. \square

Theorem 6.1. Let $n_c \leq n$, suppose there exist $Q, P, \hat{Q}, \hat{P} \in \mathbb{R}^n$ satisfying

$$0 = A\hat{Q} + \hat{Q}A^T + V_1 + \gamma^{-2}Q R_{1\infty}Q - Q\bar{E}Q + \tau_{\perp}Q\bar{E}Q\tau_{\perp}^T, \quad (6.5)$$

$$0 = (A + \gamma^{-2}[Q + \hat{Q}]R_{1\infty})^T P + P(A + \gamma^{-2}[Q + \hat{Q}]R_{1\infty}) + R_1 \\ - S^T P \Sigma P S + \tau_{\perp}^T S^T P \Sigma P S \tau_{\perp}, \quad (6.6)$$

$$0 = (A - \Sigma P S + \gamma^{-2}Q R_{1\infty})\hat{Q} + \hat{Q}(A - \Sigma P S + \gamma^{-2}Q R_{1\infty})^T \\ + \gamma^{-2}\hat{Q}(R_{1\infty} + \beta^2 S^T P \Sigma P S)\hat{Q} + Q\bar{E}Q - \tau_{\perp}Q\bar{E}Q\tau_{\perp}^T, \quad (6.7)$$

$$0 = (A - Q\bar{E} + \gamma^{-2}Q R_{1\infty})^T \hat{P} + \hat{P}(A - Q\bar{E} + \gamma^{-2}Q R_{1\infty}) \\ + S^T P \Sigma P S - \tau_{\perp}^T S^T P \Sigma P S \tau_{\perp}, \quad (6.8)$$

$$\text{rank } \hat{Q} = \text{rank } \hat{P} = \text{rank } \hat{Q}\hat{P} = n_c, \quad (6.9)$$

and let (A_c, B_c, C_c, Q) be given by

$$A_c = \Gamma(A - Q\bar{E} - \Sigma P S + \gamma^{-2}Q R_{1\infty})G^T, \quad (6.10)$$

$$B_c = \Gamma Q C^T V_2^{-1}, \quad (6.11)$$

$$C_c = -R_2^{-1}B^T P S G^T, \quad (6.12)$$

$$Q = \begin{bmatrix} Q + \hat{Q} & \hat{Q}\Gamma^T \\ \Gamma\hat{Q} & \Gamma\hat{Q}\Gamma^T \end{bmatrix}. \quad (6.13)$$

Then, $(\tilde{A}, [\gamma^{-2} Q \tilde{R}_\infty Q + \tilde{V}]^{\frac{1}{2}})$ is stabilizable if and only if \tilde{A} is asymptotically stable. In this case, the closed-loop transfer function $H(s)$ satisfies the H_∞ disturbance attenuation constraint

$$\|H(s)\|_\infty \leq \gamma, \quad (6.14)$$

and the L_2 performance criterion (2.7) satisfies the bound

$$J(A_e, B_e, C_e) \leq \text{tr}[(Q + \hat{Q})R_1 + \hat{Q}S^T P \Sigma P S]. \quad (6.15)$$

Remark 6.1. It is easy to see that Theorem 6.1 is a direct generalization of Theorem 4.1. To recover Theorem 4.1, set $n_e = n$ so that $r = G = \Gamma = I_n$ and $r_\perp = 0$. In this case the last term in each of (6.5)–(6.8) can be deleted and equation (6.8) becomes superfluous. Furthermore, (6.5)–(6.7) now reduce to (3.7)–(3.9), as expected. If, furthermore, $\beta = 0$ then $S = I_n$ so that equations (6.5)–(6.7) now reduce to the “cheap” H_∞ control case given by (3.7), (3.14) and (3.15). Alternatively, setting $\gamma = \infty$ and retaining the reduced-order constraint $n_e < n$ yields the result of [20].

Remark 6.2. Consider the case $R_{1\infty} = R_1$ and $\beta = 1$. By introducing a new variable $Z = PS = (P^{-1} + \gamma^{-2}\hat{Q})^{-1}$ as in Section 5, equation (6.6) becomes

$$\begin{aligned} 0 = & (A + \gamma^{-2}QR_1)^T Z + Z(A + \gamma^{-2}QR_1) + R_1 \\ & - Z \Sigma Z + r_\perp^T Z \Sigma Z r_\perp + \gamma^{-2}Z[Q \tilde{\Sigma} Q - r_\perp Q \tilde{\Sigma} Q r_\perp^T] Z. \end{aligned} \quad (6.16)$$

Note that (6.16) specializes to (5.7) when $r_\perp = 0$ (i.e., $n_e = n$). Furthermore, PS can be replaced by Z in (6.7)–(6.12). Next, to generalize (5.17), define P_0 as in (5.28) so that (5.17) becomes

$$\begin{aligned} 0 = & A^T P_0 + P_0 A + R_1 + \gamma^{-2} P_0 V_1 P_0 - P_0 \Sigma P_0 \\ & + (I_n - \gamma^{-2} P_0 Q) r_\perp^T (I_n - \gamma^{-2} P_0 Q)^{-1} P_0 \Sigma P_0 (I_n - \gamma^{-2} P_0 Q)^{-1} r_\perp (I_n - \gamma^{-2} P_0 Q). \end{aligned} \quad (6.17)$$

Again, (6.17) specializes to (5.17) when $r_\perp = 0$.

7. Analysis of the Design Equations

Before developing numerical algorithms, it is instructive to analyze the design equations to determine existence and multiplicity of nonnegative-definite solutions. Although a detailed theoretical analysis remains an area for future research, in this section we present a simplified treatment which highlights important asymptotic properties of the equations. It turns out that several key properties are discernible by considering the scalar case $n = 1$. Although many of these properties

can be developed for general n , the simplified scalar case suffices for obtaining a useful qualitative analysis. Here we consider only (3.7), (3.14) and (3.15).

Since the Q equation (3.7) is decoupled from (3.14) and (3.15), it can be analyzed separately. It is easy to see that there exists a unique nonnegative solution for $\gamma > (R_1/\bar{\Sigma})^{\frac{1}{2}}$ as in the case of a standard Riccati equation with stabilizability and detectability hypothesis. Furthermore, it can be seen that for

$$\left(R_1/[\bar{\Sigma} + (A^2/V_1)]\right)^{\frac{1}{2}} < \gamma < (R_1/\bar{\Sigma})^{\frac{1}{2}}$$

there exist two nonnegative solutions when A is stable and zero nonnegative solutions when A is unstable. Below this lower bound for γ nonnegative solutions Q do not exist. This result thus indicates (as in LQG theory [42]) a lower bound to the achievable H_∞ disturbance attenuation as determined by the sensor noise intensity V_2 appearing in $\bar{\Sigma}$.

Since the P and \hat{Q} equations (3.14) and (3.15) are coupled they must be analyzed jointly. Since (3.15) is a standard Riccati equation, it follows under generic hypotheses that it possesses exactly one nonnegative-definite solution for all values of Q and \hat{Q} . The analysis of the \hat{Q} equation is, however, more involved. It can be shown that the existence of real solutions is a complicated function of γ, Q , and P . When real solutions do exist, it follows that there exist either zero or two nonnegative-definite solutions. To obtain further qualitative insight into the solutions P and \hat{Q} we fix γ and allow $R_2 \rightarrow 0$, that is, the cheap L_2 control case. It thus follows that $P \sim (R_1/\Sigma)^{\frac{1}{2}}$ and that either $\hat{Q} \sim 2\gamma^2(\Sigma/R_1)^{\frac{1}{2}}$ or $\hat{Q} \sim \frac{1}{2}\bar{\Sigma}Q^2(\Sigma R_1)^{-\frac{1}{2}}$, which correspond to the previously mentioned pair of solutions satisfying (3.15). This result thus indicates that an arbitrarily small H_∞ disturbance attenuation constraint γ can be achieved (subject to the solvability of (3.7)) by sufficiently increasing the L_2 controller authority. That is, since solutions exist in the cheap L_2 control case, the H_∞ disturbance attenuation constraint is achievable. The ability to achieve small γ is also attributable to the fact that since $\beta = 0$, H_∞ disturbance attenuation to the control variables is not limited in (3.7), (3.14) and (3.15) as in Theorems 3.1 and 6.1. Of course, as is well known, it is not possible to make $\gamma \rightarrow 0$ by letting $\Sigma \rightarrow \infty$ and $\bar{\Sigma} \rightarrow \infty$ when the system possesses nonminimum phase zeros. Also, note that both of the asymptotic solutions to (3.15) are guaranteed to yield the bound (4.1). The solution of interest, however, is $\hat{Q} = O(\Sigma^{-\frac{1}{2}})$ since it clearly yields a lower value of $J(A_c, B_c, C_c, Q)$ than $\hat{Q} = O(\Sigma^{\frac{1}{2}})$. Finally, similar analysis can be applied to (5.6)–(5.8) and (5.16)–(5.18).

8. Numerical Algorithm and Illustrative Results

In this section we describe a numerical algorithm which has been developed and implemented for solving the coupled Riccati equations (3.7), (3.14) and (3.15). We also present numerical results for an illustrative example.

Coupled modified Riccati equations arise in a variety of problems and homotopic continuation methods have been shown to be particularly successful [23–25]. To solve (3.7), (3.14) and (3.15) we have implemented a simplified continuation method involving the constraint constant γ . The idea is to exploit the fact that for large γ the problem is approximated by LQG which provides a reliable starting solution. The continuation parameter γ is then successively decreased until either a desired value of γ is achieved or no further decrease is possible. This algorithm is now summarized. Let $\varepsilon > 0$ denote a convergence criterion.

Algorithm 8.1. To solve (3.7), (3.14) and (3.15), perform the following steps:

Step 1: Initialize $\gamma > 0$;

Step 2: Solve (3.7) for Q ;

Step 3: Let $k = 0, \hat{Q}_0 = 0$;

Step 4: Solve (3.14) for $P_{k+1} = P$ with $\hat{Q} = \hat{Q}_k$;

Step 5: Solve (3.15) for $\hat{Q}_{k+1} = \hat{Q}$ with $P = P_{k+1}$;

Step 6: If $k \geq 1$ check for $\|P_{k+1} - P_k\| < \varepsilon$ and $\|\hat{Q}_{k+1} - \hat{Q}_k\| < \varepsilon$;

Step 7: If convergence is not achieved in Step 6 (or $k = 0$) let $k \leftarrow k + 1$ and go to Step 4;
otherwise decrease γ and go to Step 2.

Steps 2, 4 and 5 were carried out using a standard Riccati solver [37] which proved to be reliable even when the quadratic term was indefinite or nonnegative definite. For instance, for the example considered below the term $\gamma^{-2}R_1 - \hat{\Sigma}$ was indefinite for all finite γ . The crucial step in the algorithm is the decreasing of γ in Step 7. Significant effort was devoted to providing a smooth transition to smaller values of γ without sacrificing computational efficiency. The development of more sophisticated continuation algorithms remains an area for future research.

The example considered was formulated in [38] and was considered extensively in [24,25,39] to

compare reduced-order design methods. The example is interesting since it possesses a complex pair of nonminimum phase zeros due to the fact that the physical system (coupled rotating disks) has noncollocated sensors and actuators. The plant is of eighth order and has two neutrally stable poles. The problem data are as follows:

$$n = n_c = 8, \quad m = \ell = 1, \quad q = p = 2,$$

$$A = \begin{bmatrix} -0.161 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ -6.004 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ -0.5822 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ -9.9835 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ -0.4073 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ -3.982 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \quad B = \begin{bmatrix} 0 \\ 0 \\ .0064 \\ .00235 \\ .0713 \\ 1.0002 \\ .1045 \\ .9955 \end{bmatrix} \quad C = [1 \quad 0_{1 \times 7}]$$

$$E_1 = E_{1\infty} = 10^{-3} \begin{bmatrix} 0 & 0 & 0 & 0 & .55 & 11 & 1.32 & 18 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix},$$

$$E_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \quad E_{2\infty} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \quad \beta = 0,$$

$$D_1 = [B \quad 0_{8 \times 1}], \quad D_2 = [0 \quad 1].$$

With the problem data as given, the LQG controller was found to yield a closed-loop H_∞ performance of 1.39 (i.e., 2.87 dB above unity gain). Using Algorithm 8.1 we obtained a solution for $\gamma = .52$ for a net H_∞ performance improvement of 8.7 dB (see Figure 1). Note that this result is consistent with Proposition 8.1 of [3] which implies that the maximum ratio of the H_∞ performance of the optimal L_2 controller to the H_∞ performance of the optimal H_∞ controller can be no more than twice the number of right-half-plane zeros, which for the present problem with two nonminimum phase zeros corresponds to a factor of 4 (i.e., 12 dB).

Our numerical experience revealed two interesting features. First, the loop between Steps 4 and 6 converged reliably. However, a critical value γ_{\min} of γ was invariably found below which solutions could not be computed. This value γ_{\min} appears to represent the best achievable H_∞ performance for the given L_2 weights. Second, for each value of $\gamma \geq \gamma_{\min}$ for which a solution

was computed, the actual H_∞ performance was close to this value revealing that the H_∞ bound is tight. For example, the actual worst-case attenuation of the $\gamma = .52$ design shown in Figure 1 is .511. Controller characteristics are given in Table 1. Note that in each case the L_2 performance bound is within 30% of the actual L_2 performance.

| H_∞ Attenuation Constraint γ | Actual H_∞ Attenuation $\ H(s)\ _\infty$ | L_2 Performance Bound $J(A_c, B_c, C_c, Q)$ | Actual L_2 Performance $J(A_c, B_c, C_c)$ |
|---|---|---|---|
| ∞ (LQG) | 1.39 | — | .143 |
| 2 | 1.18 | .159 | .146 |
| 1.5 | 1.06 | .171 | .151 |
| 1.0 | .855 | .204 | .168 |
| .9 | .797 | .217 | .176 |
| .8 | .732 | .236 | .187 |
| .7 | .661 | .262 | .203 |
| .52 | .511 | .299 | .262 |

Table 1

Acknowledgments. We wish to thank Professor P. P. Khargonekar for several helpful discussions and for providing a preprint of [19], J. Strachla for transforming the original manuscript into T_EX, A. Daubendiek, S. Greeley, S. Richter, and A. Tellez for developing the numerical algorithm and performing the calculations of Section 8, D. Hyland, E. Collins, and L. Davis for helpful discussions and suggestions, A. N. Madiwale for providing simplifications of (6.6)–(6.8), and the reviewers for several helpful comments.

Appendix A: Proof of Theorem 6.1

To optimize (2.20) over the open set \mathcal{X} subject to the constraint (2.14), form the Lagrangian

$$\mathcal{L}(A_c, B_c, C_c, Q, P, \lambda) \triangleq \text{tr}\{\lambda Q \tilde{R} + [\tilde{A}Q + Q\tilde{A}^T + \gamma^{-2}Q\tilde{R}_\infty Q + \tilde{V}]P\}, \quad (\text{A.1})$$

where the Lagrange multipliers $\lambda \geq 0$ and $P \in \mathbb{R}^{\tilde{n} \times \tilde{n}}$ are not both zero. We thus obtain

$$\frac{\partial \mathcal{L}}{\partial Q} = (\tilde{A} + \gamma^{-2}Q\tilde{R}_\infty)^T P + P(\tilde{A} + \gamma^{-2}Q\tilde{R}_\infty) + \lambda \tilde{R}. \quad (\text{A.2})$$

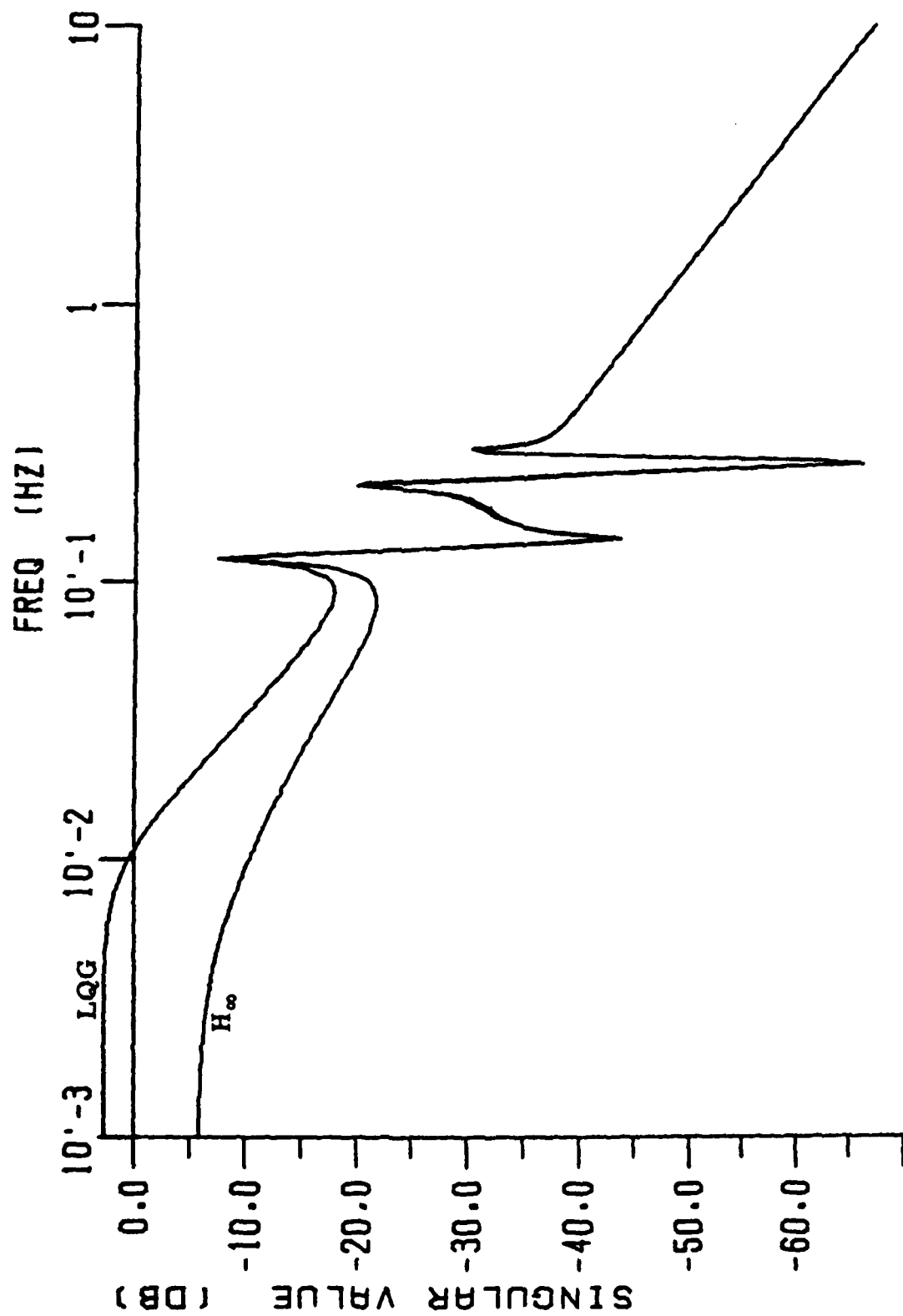


Figure 1

Setting $\partial \mathcal{L} / \partial Q = 0$ yields

$$0 = (\tilde{A} + \gamma^{-2} Q \tilde{R}_\infty)^T P + P(\tilde{A} + \gamma^{-2} Q \tilde{R}_\infty) + \lambda \tilde{R}. \quad (\text{A.3})$$

Since $\tilde{A} + \gamma^{-2} Q \tilde{R}_\infty$ is assumed to be stable, $\lambda = 0$ implies $P = 0$. Hence, it can be assumed without loss of generality that $\lambda = 1$. Furthermore, P is nonnegative definite.

Now partition $\tilde{n} \times \tilde{n}$ Q, P into $n \times n$, $n \times n_c$, and $n_c \times n_c$ subblocks as

$$Q = \begin{bmatrix} Q_1 & Q_{12} \\ Q_{12}^T & Q_2 \end{bmatrix}, \quad P = \begin{bmatrix} P_1 & P_{12} \\ P_{12}^T & P_2 \end{bmatrix}.$$

Thus, with $\lambda = 1$ the stationarity conditions are given by

$$\frac{\partial \mathcal{L}}{\partial Q} = (\tilde{A} + \gamma^{-2} Q \tilde{R}_\infty)^T P + P(\tilde{A} + \gamma^{-2} Q \tilde{R}_\infty) + \tilde{R} = 0, \quad (\text{A.4})$$

$$\frac{\partial \mathcal{L}}{\partial A_c} = P_{12}^T Q_{12} + P_2 Q_2 = 0, \quad (\text{A.5})$$

$$\frac{\partial \mathcal{L}}{\partial B_c} = P_2 B_c V_2 + (P_{12}^T Q_1 + P_2 Q_{12}^T) C^T = 0, \quad (\text{A.6})$$

$$\frac{\partial \mathcal{L}}{\partial C_c} = R_2 C_c Q_2 + \beta^2 \gamma^{-2} R_2 C_c (P_1^T Q_{12} + P_{12} Q_2)^T Q_{12} + B^T (P_1 Q_{12} + P_{12} Q_2) = 0. \quad (\text{A.7})$$

Expanding (2.14) and (A.4) yields

$$0 = A Q_1 + Q_1 A^T + B C_c Q_{12}^T + Q_{12} C_c^T B^T + \gamma^{-2} Q_1 R_{1\infty} Q_1 + \beta^2 \gamma^{-2} Q_{12} C_c^T R_2 C_c Q_{12}^T + V_1, \quad (\text{A.8})$$

$$0 = A Q_{12} + Q_{12} A_c^T + B C_c Q_2 + Q_1 C^T B_c^T + \gamma^{-2} Q_1 R_{1\infty} Q_{12} + \beta^2 \gamma^{-2} Q_{12} C_c^T R_2 C_c Q_2, \quad (\text{A.9})$$

$$0 = A_c Q_2 + Q_2 A_c^T + B_c C Q_{12} + Q_{12}^T C^T B^T + \gamma^{-2} Q_{12}^T R_{1\infty} Q_{12} + \beta^2 \gamma^{-2} Q_2 C_c^T R_2 C_c Q_2 + B_c V_2 B_c^T, \quad (\text{A.10})$$

$$0 = A^T P_1 + P_1 A + C^T B_c^T P_{12}^T + P_{12} B_c C + \gamma^{-2} R_{1\infty} (P_1 Q_1 + P_{12} Q_{12}^T)^T + \gamma^{-2} (P_1 Q_1 + P_{12} Q_{12}^T) R_{1\infty} + R_1, \quad (\text{A.11})$$

$$0 = A^T P_{12} + P_{12} A_c + C^T B_c^T P_2 + P_1 B C_c + \gamma^{-2} R_{1\infty} (P_{12}^T Q_1 + P_2 Q_{12}^T)^T + \beta^2 \gamma^{-2} (P_1 Q_{12} + P_{12} Q_2) C_c^T R_2 C_c, \quad (\text{A.12})$$

$$0 = A_c^T P_2 + P_2 A_c + P_{12}^T B C_c + C_c^T B^T P_{12} + C_c^T R_2 C_c. \quad (\text{A.13})$$

Lemma A.1. Q_2 and P_2 are positive definite.

Proof. By a minor extension of results from [40], (A.10) can be rewritten as

$$0 = (A_c + B_c C Q_{12} Q_2^+) Q_2 + Q_2 (A_c + B_c C Q_{12} Q_2^+)^T + \Psi,$$

where

$$\Psi \triangleq \gamma^{-2} Q_{12}^T R_{1\infty} Q_{12} + \beta^2 \gamma^{-2} Q_2 C_c^T R_2 C_c Q_2 + B_c V_2 B_c^T$$

and Q_2^+ is the Moore-Penrose or Drazin generalized inverse of Q_2 . Next note that since (A_c, B_c) is controllable it follows from Lemma 2.1 and Theorem 3.6 of [28] that $(A_c + B_c C Q_{12} Q_2^+, \Psi^{\frac{1}{2}})$ is also controllable. Now, since Q_2 and Ψ are nonnegative definite, Lemma 12.2 of [28] implies that Q_2 is positive definite. Using (A.13), similar arguments show that P_2 is positive definite. \square

Since R_2, V_2, Q_2, P_2 are invertible, (A.5)–(A.7) can be written as

$$-P_2^{-1} P_{12}^T Q_{12} Q_2^{-1} = I_{n_c}, \quad (\text{A.14})$$

$$B_c = -P_2^{-1} (P_{12}^T Q_1 + P_2 Q_{12}^T) C^T V_2^{-1}, \quad (\text{A.15})$$

$$C_c [I_{n_c} + \beta^2 \gamma^{-2} (Q_{12}^T P_1 + Q_2 P_{12}^T) Q_{12} Q_2^{-1}] = -R_2^{-1} B^T (P_1 Q_{12} + P_{12} Q_2) Q_2^{-1}. \quad (\text{A.16})$$

Now define the $n \times n$ matrices

$$Q \triangleq Q_1 - Q_{12} Q_2^{-1} Q_{12}^T, \quad P \triangleq P_1 - P_{12} P_2^{-1} P_{12}^T,$$

$$\hat{Q} \triangleq Q_{12} Q_2^{-1} Q_{12}^T, \quad \hat{P} \triangleq P_{12} P_2^{-1} P_{12}^T,$$

$$r \triangleq -Q_{12} Q_2^{-1} P_2^{-1} P_{12}^T,$$

and the $n_c \times n$, $n_c \times n_c$ and $n_c \times n$ matrices

$$G \triangleq Q_2^{-1} Q_{12}^T, \quad M \triangleq Q_2 P_2, \quad \Gamma \triangleq -P_2^{-1} P_{12}^T.$$

Note that $r = G^T \Gamma$.

Clearly, Q, P, \hat{Q} and \hat{P} are symmetric and \hat{Q} and \hat{P} are nonnegative definite. To show that Q and P are also nonnegative definite, note that Q is the upper left-hand block of the nonnegative-definite matrix $\tilde{Q} \tilde{Q} \tilde{Q}^T$, where

$$\tilde{Q} \triangleq \begin{bmatrix} I_n & -Q_{12} Q_2^{-1} \\ 0_{n_c \times n} & I_{n_c} \end{bmatrix}.$$

Similarly, P is nonnegative definite.

Next note that with the above definitions (A.14) is equivalent to (6.2) and that (6.1) holds. Hence $\tau = G^T \Gamma$ is idempotent, i.e., $\tau^2 = \tau$.

It is helpful to note the identities

$$\hat{Q} = Q_{12}G = G^T Q_{12}^T = G^T Q_2 G, \quad \hat{P} = -P_{12}\Gamma = -\Gamma^T P_{12}^T = \Gamma^T P_2 \Gamma, \quad (\text{A.17})$$

$$\tau G^T = G^T, \quad \Gamma \tau = \Gamma, \quad (\text{A.18})$$

$$\hat{Q} = \tau \hat{Q}, \quad \hat{P} = \hat{P} \tau, \quad (\text{A.19})$$

$$\hat{Q} \hat{P} = -Q_{12} P_{12}^T. \quad (\text{A.20})$$

Using (A.14) and Sylvester's inequality, it follows that

$$\text{rank } G = \text{rank } \Gamma = \text{rank } Q_{12} = \text{rank } P_{12} = n_c.$$

Now using (A.17) and Sylvester's inequality yields

$$n_c = \text{rank } Q_{12} + \text{rank } G - n_c \leq \text{rank } \hat{Q} \leq \text{rank } Q_{12} = n_c,$$

which implies that $\text{rank } \hat{Q} = n_c$. Similarly, $\text{rank } \hat{P} = n_c$, and $\text{rank } \hat{Q} \hat{P} = n_c$ follows from (A.20).

The components of Q and P can be written in terms of $Q, P, \hat{Q}, \hat{P}, G$ and Γ as

$$Q_1 = Q + \hat{Q}, \quad P_1 = P + \hat{P}, \quad (\text{A.21})$$

$$Q_{12} = \hat{Q} \Gamma^T, \quad P_{12} = -\hat{P} G^T, \quad (\text{A.22})$$

$$Q_2 = \Gamma \hat{Q} \Gamma^T, \quad P_2 = G \hat{P} G^T. \quad (\text{A.23})$$

Next note that by using (A.21)–(A.23) it can be shown that the right-hand coefficient of C_c in (A.16) is given by

$$\hat{S} \triangleq I_{n_c} + \beta^2 \gamma^{-2} \Gamma \hat{Q} P G^T.$$

To prove that \hat{S} is invertible use (A.19) and (6.3) and note that

$$\begin{aligned} I_{n_c} + \beta^2 \gamma^{-2} \Gamma \hat{Q} P G^T &= I_{n_c} + \beta^2 \gamma^{-2} \Gamma \hat{Q} \tau^T P G^T \\ &= I_{n_c} + \beta^2 \gamma^{-2} (\Gamma \hat{Q} \Gamma^T) (G P G^T). \end{aligned}$$

Since $\Gamma \hat{Q} \Gamma^T$ and $G P G^T$ are nonnegative definite, their product has nonnegative eigenvalues (see Lemma 5.1). Thus each eigenvalue of $I_{n_c} + \beta^2 \gamma^{-2} \Gamma \hat{Q} P G^T$ is real and is greater than unity. Hence \hat{S} is invertible. Now note that by using (6.2) and (6.3) it can be shown that

$$G^T \hat{S}^{-1} \Gamma = S \tau.$$

The expressions (6.11), (6.12) and (6.13) follow from (A.15), (A.16), and the definition of \mathcal{Q} . Next, computing either $\Gamma(A.9)-(A.10)$ or $G(A.12)+(A.13)$ yields (6.10). Substituting (A.21)-(A.23) into (A.8)-(A.13) and the expression for A_c into (A.9), (A.10), (A.12) and (A.13) it follows that (A.10) = $\Gamma(A.9)$ and (A.13) = $G(A.12)$. Thus, (A.10) and (A.13) are superfluous and can be omitted. Thus, (A.8)-(A.13) reduce to

$$0 = AQ + QA^T + V_1 + \gamma^{-2}(Q + \hat{Q})R_{1\infty}(Q + \hat{Q}) + \beta^2\gamma^{-2}\hat{Q}S^TP\Sigma PS\hat{Q} \\ + (A - \Sigma PS)\hat{Q} + \hat{Q}(A - \Sigma PS)^T, \quad (A.24)$$

$$0 = [(A - \Sigma PS)\hat{Q} + \hat{Q}(A - \Sigma PS)^T + Q\bar{\Sigma}Q + \gamma^{-2}(Q + \hat{Q})R_{1\infty}(Q + \hat{Q}) - \gamma^{-2}QR_{1\infty}Q \\ + \beta^2\gamma^{-2}\hat{Q}S^TP\Sigma PS\hat{Q}]\Gamma^T, \quad (A.25)$$

$$0 = (A + \gamma^{-2}[Q + \hat{Q}]R_{1\infty})^TP + P(A + \gamma^{-2}[Q + \hat{Q}]R_{1\infty}) + R_1 \\ + (A - Q\bar{\Sigma} + \gamma^{-2}QR_{1\infty})^T\hat{P} + \hat{P}(A - Q\bar{\Sigma} + \gamma^{-2}QR_{1\infty}), \quad (A.26)$$

$$0 = [(A - Q\bar{\Sigma} + \gamma^{-2}QR_{1\infty})^T\hat{P} + \hat{P}(A - Q\bar{\Sigma} + \gamma^{-2}QR_{1\infty}) + S^TP\Sigma PS]G^T. \quad (A.27)$$

Next, using (A.24)+ $G^T\Gamma(A.25)G-(A.25)G-[(A.25)G]^T$ and $G^T\Gamma(A.25)G-(A.25)G-[(A.25)G]^T$ yields (6.5) and (6.7). Similarly, using (A.26)+ $\Gamma^TG(A.27)\Gamma-(A.27)\Gamma-[(A.27)\Gamma]^T$ and $\Gamma^TG(A.27)\Gamma-(A.27)\Gamma-[(A.27)\Gamma]^T$ yields (6.6) and (6.8).

Finally, to prove the converse we use (6.5)-(6.13) to obtain (2.14) and (A.4)-(A.7). Let $A_c, B_c, C_c, G, \Gamma, \tau, Q, P, \hat{Q}, \hat{P}, \mathcal{Q}$ be as in the statement of Theorem 6.1 and define $Q_1, Q_{12}, Q_2, P_1, P_{12}, P_2$ by (A.21)-(A.23). Using (6.2), (6.11) and (6.12) it is easy to verify (A.6) and (A.7). Finally, substitute the definitions of $Q, P, \hat{Q}, \hat{P}, G, \Gamma$, and τ into (6.5)-(6.8) using (6.2), (6.3), and (A.19) to obtain (2.14) and (A.4). Finally, note that

$$\mathcal{Q} = \begin{bmatrix} Q & 0_{n \times n_c} \\ 0_{n_c \times n} & 0_{n_c} \end{bmatrix} + \begin{bmatrix} I_n \\ \Gamma \end{bmatrix} \hat{Q} [I_n \quad \Gamma^T],$$

which shows that $\mathcal{Q} \geq 0$. \square

References

1. H. Kwakernaak and R. Sivan, *Linear Optimal Control Systems*, Wiley, New York, 1972.
2. G. Zames, "Feedback and Optimal Sensitivity: Model Reference Transformations, Multiplicative Seminorms, and Approximate Inverses," *IEEE Trans. Autom. Contr.*, Vol. AC-26, pp. 301-320, 1981.
3. B. A. Francis and J. C. Doyle, "Linear Control Theory with an H_∞ Optimality Criterion," *SIAM J. Contr. Optim.*, Vol. 25, pp. 815-844, 1987.
4. B. A. Francis, *A Course in H_∞ Control Theory*, Springer-Verlag, New York, 1987.
5. J. C. Doyle and G. Stein, "Multivariable Feedback Design: Concepts for a Classical/Modern Synthesis," *IEEE Trans. Autom. Contr.*, Vol. AC-26, pp. 4-16, 1981.
6. B. A. Francis, "On Robustness of the Stability of Feedback Systems," *IEEE Trans. Autom. Contr.*, Vol. AC-25, pp. 817-818, 1980.
7. I. R. Petersen, "A Riccati Equation Approach to the Design of Stabilizing Controllers and Observers for a Class of Uncertain Linear Systems," *IEEE Trans. Autom. Contr.*, Vol. AC-30, pp. 904-907, 1985.
8. A. R. Galimidi and B. R. Barmish, "The Constrained Lyapunov Problem and Its Application to Robust Output Feedback Stabilization," *IEEE Trans. Autom. Contr.*, Vol. AC-31, pp. 410-419, 1986.
9. I. R. Petersen and C. V. Hollot, "A Riccati Equation Approach to the Stabilization of Uncertain Systems," *Automatica*, Vol. 22, pp. 397-411, 1986.
10. K. Zhou and P. P. Khargonekar, "Stability Robustness Bounds for Linear State-Space Models with Structured Uncertainty," *IEEE Trans. Autom. Contr.*, Vol. AC-32, pp. 621-623, 1987.
11. D. Hinrichsen and A. J. Pritchard, "Stability Radius for Structured Perturbations and the Algebraic Riccati equation," *Sys. Contr. Lett.*, Vol. 8, pp. 105-113, 1987.
12. D. S. Bernstein and W. M. Haddad, "The Optimal Projection Equations with Petersen-Hollot Bounds: Robust Controller Synthesis with Guaranteed Structured Stability Radius," *Proc. IEEE Conf. Dec. Contr.*, pp. 1308-1318, Los Angeles, CA, December 1987.
13. D. S. Bernstein, "Robust Static and Dynamic Output-Feedback Stabilization: Deterministic and Stochastic Perspectives," *IEEE Trans. Autom. Contr.*, Vol. AC-32, pp. 1076-1084, 1987.
14. D. S. Bernstein and W. M. Haddad, "The Optimal Projection Equations with Petersen-Hollot Bounds: Robust Stability and Performance via Fixed-Order Dynamic Compensation for Systems with Structured Real-Valued Parameter Uncertainty," *IEEE Trans. Autom. Contr.*, Vol. AC-33, pp. 578-582, 1988.
15. D. S. Bernstein and W. M. Haddad, "Robust Stability and Performance for Fixed-Order Dynamic Compensation via the Optimal Projection Equations with Guaranteed Cost Bounds," *Proc. Amer. Contr. Conf.*, pp. 2471-2476, Atlanta, GA, June 1988.
16. D. S. Bernstein and W. M. Haddad "Robust Stability and Performance Analysis for Linear Dynamic Systems," submitted.
17. J. C. Doyle, "Analysis of Feedback Systems with Structured Uncertainties," *IEE Proc.*, Vol. 129, pp. 242-250, 1982.
18. I. R. Petersen, "Disturbance Attenuation and H^∞ Optimization: A Design Method Based on the Algebraic Riccati equation," *IEEE Trans. Autom. Contr.*, Vol. AC-32, pp. 427-429, 1987.
19. P. P. Khargonekar, I. R. Petersen, and K. Zhou, "Robust Stabilization of Uncertain Systems and H^∞ Optimal Control," preprint.

20. D. C. Hyland and D. S. Bernstein, "The Optimal Projection Equations for Fixed-Order Dynamic Compensation," *IEEE Trans. Autom. Contr.*, Vol. AC-29, pp. 1034-1037, 1984.
21. D. A. Wilson, "Convolution and Hankel Operator Norms for Linear Systems," preprint.
22. M. J. Grimble, "Optimal H_∞ Robustness and the Relationship to LQG Design Problems," *Int. J. Contr.*, Vol. 43, pp. 351-372, 1986.
23. M. Mariton and R. Bertrand, "A Homotopy Algorithm for Solving Coupled Riccati Equations," *Optim. Contr. Appl. Meth.*, Vol. 6, pp. 351-357, 1985.
24. S. Richter, "A Homotopy Algorithm for Solving the Optimal Projection Equations for Fixed-Order Dynamic Compensation: Existence, Convergence and Global Optimality," *Proc. Amer. Contr. Conf.*, pp. 1527-1531, Minneapolis, MN, June 1987.
25. S. W. Greeley and D. C. Hyland, "Reduced-Order Compensation: LQG Reduction Versus Optimal Projection Using a Homotopic Continuation Method," *Proc. IEEE Conf. Dec. Contr.*, pp. 742-747, Los Angeles, CA, December 1987.
26. J. C. Doyle, K. Glover, P. P. Khargonekar, and B. A. Francis, "State-Space Solutions to Standard H_2 and H_∞ Control Problems," *Proc. Amer. Contr. Conf.*, pp. 1691-1696, Atlanta, GA, June 1988.
27. A. N. Madiwale, "Some Practical Aspects of H_∞ Compensators," MIT Lincoln Laboratory, in preparation.
28. W. M. Wonham, *Linear Multivariable Control: A Geometric Approach*, Springer-Verlag, New York, 1979.
29. J. C. Willems, "Least Squares Stationary Optimal Control and the Algebraic Riccati Equation," *IEEE Trans. Autom. Contr.*, Vol. AC-16, pp. 621-634, 1971.
30. R. W. Brockett, *Finite Dimensional Linear Systems*, Wiley, New York, 1970.
31. K. Glover, "All Optimal Hankel-Norm Approximations of Linear Multivariable Systems and Their L^∞ -Error Bounds," *Int. J. Contr.*, Vol. 39, pp. 1115-1193, 1984.
32. D. S. Bernstein, "The Optimal Projection Equations for Static and Dynamic Output Feedback: The Singular Case," *IEEE Trans. Autom. Contr.*, Vol. AC-32, pp. 1139-1143, 1987.
33. A. Bensoussan and J. H. van Schuppen, "Optimal Control of Partially Observable Stochastic Systems With an Exponential-of-Integral Performance Index," *SIAM J. Contr. Optim.*, Vol. 23, pp. 599-613, 1985.
34. D. H. Jacobson, *Extensions of Linear-Quadratic Control, Optimization and Matrix Theory*, Academic Press, London, 1977.
35. P. Whittle, "Risk-Sensitive Linear/Quadratic/Gaussian Control," *Adv. Appl. Prob.*, Vol. 13, pp. 764-777, 1981.
36. C. R. Rao and S. K. Mitra, *Generalized Inverse of Matrices and Its Applications*, John Wiley and Sons, New York, 1971.
37. A. J. Laub, "A Schur Method for Solving Riccati Equations," *IEEE Trans. Autom. Contr.*, Vol. AC-32, pp. 913-921, 1979.
38. R. H. Cannon, Jr., and D. E. Rosenthal, "Experiments in Control of Flexible Structures with Noncolocated Sensors and Actuators," *AIAA J. Guid. Contr. Dyn.*, Vol. 7, pp. 546-553, 1984.
39. Y. Liu and B. D. O. Anderson, "Controller Reduction via Stable Factorization and Balancing," *Int. J. Contr.*, Vol. 44, pp. 507-531, 1986.
40. A. Albert, "Conditions for Positive and Nonnegative Definiteness in Terms of Pseudo Inverse," *SIAM J. Contr. Optim.*, Vol. 17, pp. 434-440, 1969.

41. D. S. Bernstein and W. M. Haddad, "LQG Control with an H_∞ Performance Bound: A Riccati Equation Approach," *Proc. Amer. Contr. Conf.*, pp. 796-802, Atlanta, GA, June 1988.
42. H. Kwakernaak and R. Sivan, "The Maximally Achievable Accuracy of Linear Optimal Regulators and Linear Optimal Filters," *IEEE Trans. Autom. Contr.*, Vol. AC-17, pp. 79-86, 1972.

ROBUST H_∞ CONTROL DESIGN FOR SYSTEMS WITH STRUCTURED PARAMETER UNCERTAINTY*

Appasaheb N. Madiwale
Lincoln Laboratory/M.I.T.
P.O. Box 73
Lexington, MA 02173

Wassim M. Haddad**
Florida Institute of Technology
Melbourne, FL 32901

Dennis S. Bernstein**
Harris Corporation
P.O. Box 37
Melbourne, FL 32901

Abstract

In a recent paper a unification of the L_2 (LQG) and H_∞ control design problems was obtained in terms of modified algebraic Riccati equations. In the present paper these results are extended to guarantee robust L_2 and H_∞ performance in the presence of structured real-valued parameter variations (ΔA , ΔB , ΔC) in the state space model. For design flexibility the paper considers two distinct types of uncertainty bounds for both full- and reduced-order dynamic compensation. An important special case of these results generates L_2/H_∞ controller designs with guaranteed gain margins.

I. Introduction

It has recently been shown that the solution to the optimal H_∞ disturbance attenuation problem can be expressed in terms of a pair of modified Riccati equations ([1,2]). Furthermore, it was shown in [1] that L_2/H_∞ design tradeoffs can be achieved by solving a coupled system consisting of three modified Riccati equations. As is well known, the disturbance attenuation problem can be used to guarantee robustness with respect to unstructured plant uncertainties. However, if plant uncertainty is present in the form of parametric variations of the state space model, then alternative bounding techniques are required. The goal of the present paper is thus to extend the results of [1] to include bounds on the effect of real-valued structured parameter variations.

In the absence of an H_∞ design constraint, robust stability and L_2 performance for dynamic compensator design were guaranteed in [3,4] by incorporating quadratic Lyapunov bounds within LQG design theory. Two distinct bounds were considered. In [3] a quadratic bound was used while in [4] a linear bound was employed. In each case full- and reduced-order dynamic compensators were characterized by means of coupled systems of modified Riccati and Lyapunov equations.

To design H_∞ controllers which are robust with respect to structured real-valued parameter variations we proceed by combining the results of [1] with those of [3,4]. That is, we derive coupled systems of modified Riccati and Lyapunov equations whose solutions yield controllers which are guaranteed to satisfy a prespecified H_∞ attenuation constraint for all variations (ΔA , ΔB , ΔC) belonging to a given uncertainty set. If the uncertainty is absent (i.e., $\Delta A = 0$, etc.) then the results of [1] are recovered, while if the H_∞ constraint is relaxed, then the results of [3,4] are obtained. Thus the results of [1] can be viewed as a specialization of a broader design theory which accounts for structured real-valued parameter uncertainty. Finally, we state all results for the case of a fixed-order (i.e., reduced-order) controller for maximal design

flexibility. Extensions to even more general design problems are mentioned in Section 9 but omitted here for lack of space.

Notation. Note: All matrices have real entries.

$R, R^{r \times s}, R^r, E$ real numbers, $r \times s$ real matrices, $R^{r \times 1}$, expected value

$I_r, ()^T, 0_{r \times s}, 0_r$ $r \times r$ identity matrix, transpose, $r \times s$ zero matrix, $0_{r \times r}$

S^r, N^r, P^r $r \times r$ symmetric, nonnegative definite, positive-definite matrices

$Z_1 \leq Z_2, Z_1 < Z_2$ $Z_2 - Z_1 \in N^r, Z_2 - Z_1 \in P^r, Z_1; Z_2 \in S^r$

n, m, ℓ, n_s positive integers

$p, d, \infty, q, q_\infty; \bar{n}$ positive integers; $n + n_c$ ($n_c \leq n$)

x, u, y, x_c, \bar{x} n, m, ℓ, n_c, \bar{n} -dimensional vectors

$A, \Delta A; B, \Delta B; C, \Delta C$ $n \times n; n \times m; \ell \times n$ matrices

A_c, B_c, C_c $n_c \times n_c, n_c \times \ell, m \times n_c$ matrices

$\bar{x}, \bar{A}, \bar{\Delta A}$ $\begin{bmatrix} x \\ x_c \end{bmatrix}, \begin{bmatrix} A & BC_c \\ B_c C & A_c \end{bmatrix}, \begin{bmatrix} \Delta A & \Delta BC_c \\ B_c \Delta C & 0 \end{bmatrix}$

$w(\cdot)$ d -dimensional standard white noise

D_1, D_2 $n \times d, \ell \times d$ matrices; $D_1 D_2^T = 0$

V_1, V_2 $D_1 D_1^T, D_2 D_2^T; V_2 \in \mathbb{P}^\ell$

\bar{D}, \bar{V} $\begin{bmatrix} D_1 \\ B_c D_2 \end{bmatrix}, \begin{bmatrix} V_1 & 0_{n \times n_c} \\ 0_{n_c \times n} & B_c V_2 B_c^T \end{bmatrix}$

E_1, E_2 $q \times n, q \times m$ matrices; $E_1^T E_2 = 0$

\bar{E}, R_1, R_2 $[E_1 \ E_2 C_c], E_1^T E_1, E_2^T E_2; R_2 \in \mathbb{P}^m$

\bar{R} $\begin{bmatrix} R_1 & 0_{n \times n_c} \\ 0_{n_c \times n} & C_c^T R_2 C_c \end{bmatrix} = \bar{E}^T \bar{E}$

$E_{1\infty}, E_{2\infty}$ $q_\infty \times n, q_\infty \times m$ matrices;
 $E_{1\infty}^T E_{2\infty} = 0$

$\bar{E}_\infty, R_{1\infty}, R_{2\infty}$ $[E_{1\infty} \ E_{2\infty} C_c], E_{1\infty}^T E_{1\infty}, E_{2\infty}^T E_{2\infty}$

\bar{R}_∞ $\begin{bmatrix} R_{1\infty} & 0_{n \times n_c} \\ 0_{n_c \times n} & C_c^T R_{2\infty} C_c \end{bmatrix} = \bar{E}_\infty^T \bar{E}_\infty$

* This work is sponsored by the Dept. of the Air Force. The views expressed are those of the author and do not reflect the official policy or position of the U.S. Government.

** Supported in part by the Air Force Office of Scientific Research under contract F49620-86-C-0002.

$D_{1\infty}, D_{2\infty}$ $n \times d_{\infty}, \ell \times d_{\infty}$ matrices; $D_{1\infty}D_{2\infty}^T = 0$

$V_{1\infty}, V_{2\infty}$ $D_{1\infty}D_{1\infty}^T, D_{2\infty}D_{2\infty}^T$

$\bar{D}_{\infty}, \bar{V}_{\infty}$ $\begin{bmatrix} D_{1\infty} \\ B_c D_{2\infty} \end{bmatrix}, \begin{bmatrix} V_{1\infty} & 0_{n \times n_c} \\ 0_{n_c \times n} & B_c V_{2\infty} B_c^T \end{bmatrix}$

β, γ, α nonnegative constant; positive constants

$A_{\alpha}, A_{c\alpha}$ $A + \frac{\alpha}{2} I_n, A_c + \frac{\alpha}{2} I_{n_c}$

II. Robust Stability and L_2 Performance with a Robust H_{∞} Constraint

In this section we state the Robust stability and L_2 performance problem with robust H_{∞} disturbance attenuation constraint. Specifically, we consider a fixed-order dynamic output-feedback control-design problem with structured real-valued plant parameter uncertainties and constrained H_{∞} disturbance attenuation. This problem involves a set $U \subset \mathbb{R}^{n \times n} \times \mathbb{R}^{n \times m} \times \mathbb{R}^{\ell \times m}$ of uncertain perturbations $(\Delta A, \Delta B, \Delta C)$ of the nominal system matrices (A, B, C) . The goal of the problem is to determine a fixed-order, strictly proper dynamic compensator (A_c, B_c, C_c) which (a) stabilizes the plant for all variations in U , (b) satisfies an H_{∞} constraint on disturbance rejection for all variations in U , and (c) minimizes the worst-case value over the uncertainty set U of a steady-state L_2 performance criterion. In this and the following section no explicit structure is assumed for the elements of U . In Sections 4 and 7, two specific structures of variations in U will be introduced.

H_{∞} - Constrained Robust Dynamic Compensation Problem. Given the n th-order stabilizable and detectable plant with structured real-valued plant parameter variations

$$\dot{x}(t) = (A + \Delta A)x(t) + (B + \Delta B)u(t) + D_1 w(t), \quad (2.1)$$

$$y(t) = (C + \Delta C)x(t) + D_2 w(t), \quad (2.2)$$

determine an n_c th-order dynamic compensator

$$\dot{x}_c(t) = A_c x_c(t) + B_c y(t) \quad (2.3)$$

$$u(t) = C_c x_c(t), \quad (2.4)$$

which satisfies the following design criteria:

(i) the closed-loop system (2.1) - (2.4) is asymptotically stable for all $(\Delta A, \Delta B, \Delta C) \in U$, i.e., $\bar{A} + \Delta \bar{A}$ is asymptotically stable for all $(\Delta A, \Delta B, \Delta C) \in U$;

(ii) the $q_{\infty} \times d$ closed-loop transfer function

$$\Pi_{\Delta \bar{A}}(s) \triangleq \bar{E}_{\infty} [sI_n - (\bar{A} + \Delta \bar{A})]^{-1} \bar{D}, \quad (2.5)$$

from $w(t)$ to $E_1 x(t) + E_2 x_c(t)$ satisfies the constraint

$$\|\Pi_{\Delta \bar{A}}(s)\|_{\infty} \leq \gamma, (\Delta A, \Delta B, \Delta C) \in U, \quad (2.6)$$

where $\gamma > 0$ is a given constant; and

(iii) the performance functional

$$J(A_c, B_c, C_c) \triangleq \sup_{(\Delta A, \Delta B, \Delta C) \in U} \limsup_{t \rightarrow \infty} E[x^T(t)R_1 x(t) + u^T(t)R_2 u(t)] \quad (2.7)$$

is minimized.

Note that for each uncertain variation $(\Delta A, \Delta B, \Delta C) \in U$, the closed-loop system can be written as

$$\dot{\bar{x}}(t) = (\bar{A} + \Delta \bar{A})\bar{x}(t) + \bar{D}w(t), t \in [0, \infty), \quad (2.8)$$

and that (2.7) becomes

$$J(A_c, B_c, C_c) = \sup_{(\Delta A, \Delta B, \Delta C) \in U} \limsup_{t \rightarrow \infty} E[\bar{x}^T(t)\bar{R}\bar{x}(t)]. \quad (2.9)$$

Note that the problem statement involves both L_2 and H_{∞} performance weights. In particular, the matrices R_1 and R_2 are the L_2 weights for the state and control variables. By introducing the variables

$$z(t) = E_1 x(t), v(t) = E_2 u(t),$$

the cost (2.7) can be written as

$$J(A_c, B_c, C_c) = \sup_{(\Delta A, \Delta B, \Delta C) \in U} \limsup_{t \rightarrow \infty} E[z^T(t)z(t) + v^T(t)v(t)].$$

For convenience we thus define $R_1 \triangleq E_1^T E_1$ and $R_2 \triangleq E_2^T E_2$ which appear in subsequent expressions. Although an L_2 cross-weighting term of the form $2x^T(t)R_{12}u(t)$ can also be included, we shall not do so here to facilitate the presentation.

For the H_{∞} performance constraint, the transfer function (2.5) involves weighting matrices $E_{1\infty}$ and $E_{2\infty}$ for the state and control variables. The matrices $R_{1\infty} \triangleq E_{1\infty}^T E_{1\infty}$ and $R_{2\infty} \triangleq E_{2\infty}^T E_{2\infty}$ are thus the H_{∞} counterparts of the L_2 weights R_1 and R_2 . Although we do not require that $R_{1\infty}$ and $R_{2\infty}$ be equal to R_1 and R_2 , we shall require that $R_{2\infty} = \beta^2 R_2$, where the nonnegative scalar β is a design variable. We further note that the assumption $E_{1\infty} E_{2\infty}^T = 0$ precludes an H_{∞} cross-weighting term which again facilitates the presentation. Finally, similar remarks apply to the disturbance and sensor noise intensities $V_1 \triangleq D_1 D_1^T$, $V_2 \triangleq D_2 D_2^T$, $V_{1\infty} \triangleq D_{1\infty} D_{1\infty}^T$ and $V_{2\infty} \triangleq D_{2\infty} D_{2\infty}^T$ for the L_2 and H_{∞} designs respectively. We note that $w(t)$ can be interpreted as white noise for the L_2 design aspect and as an L_2 signal for the H_{∞} design aspect [1].

Before continuing it is useful to note that if $(\bar{A} + \Delta \bar{A})$ is asymptotically stable for all $(\Delta A, \Delta B, \Delta C) \in U$ for a given compensator (A_c, B_c, C_c) then the performance (2.7) is given by

$$J(A_c, B_c, C_c) = \sup_{(\Delta A, \Delta B, \Delta C) \in U} \text{tr } \bar{Q}_{\Delta A} \bar{R}, \quad (2.10)$$

where the steady-state closed-loop state covariance defined by

$$\bar{Q}_{\Delta A} \triangleq \lim_{t \rightarrow \infty} E[\bar{x}(t) \bar{x}(t)^T], \quad (2.11)$$

satisfies the $\bar{n} \times \bar{n}$ algebraic Lyapunov equation

$$0 = (\bar{A} + \Delta \bar{A}) \bar{Q}_{\Delta A} + \bar{Q}_{\Delta A} (\bar{A} + \Delta \bar{A})^T + \bar{V}. \quad (2.12)$$

The key step in guaranteeing robust stability and performance is to replace the uncertain terms in the covariance Lyapunov equation (2.12) by a bounding function Ω . Note that since $\Delta \bar{A}$ is independent of A_c , the bounding function Ω need only depend on B_c and C_c . Furthermore, the H_∞ constraint (2.6) on the disturbance attenuation is enforced for all $(\Delta A, \Delta B, \Delta C) \in U$ by replacing the modified algebraic Lyapunov equation (2.12) by an algebraic Riccati equation which overbounds the closed-loop steady-state covariance. Justification for this technique is provided by the following result.

Lemma 2.1. Let $\Omega: \mathbb{R}^{n_c \times \ell} \times \mathbb{R}^{m \times n_c} \times \mathbb{N}^{\bar{n}} \rightarrow \mathbb{S}^{\bar{n}}$ be such that

$$\Delta \bar{A} \bar{Q} + \bar{Q} \Delta \bar{A}^T \leq \Omega(B_c, C_c, Q); (\Delta A, \Delta B, \Delta C) \in U, \\ (B_c, C_c, Q) \in \mathbb{R}^{n_c \times \ell} \times \mathbb{R}^{m \times n_c} \times \mathbb{N}^{\bar{n}}, \quad (2.13)$$

and, for a given (A_c, B_c, C_c) , suppose there exists $Q \in \mathbb{N}^{\bar{n}}$ satisfying

$$0 = \bar{A} \bar{Q} + \bar{Q} \bar{A}^T + \gamma^{-2} \bar{Q} \bar{R} \bar{Q} + \Omega(B_c, C_c, Q) + \bar{V}, \quad (2.14)$$

then

$$(\bar{A} + \Delta \bar{A}, [\bar{V} + \gamma^{-2} \bar{Q} \bar{R} \bar{Q} + \Omega(B_c, C_c, Q) - (\Delta \bar{A} \bar{Q} + \bar{Q} \Delta \bar{A}^T)]^{\frac{1}{2}}) \\ \text{is stabilizable for all } (\Delta A, \Delta B, \Delta C) \in U. \quad (2.15)$$

if and only if

$$\bar{A} + \Delta \bar{A} \text{ is asymptotically stable, } (\Delta A, \Delta B, \Delta C) \in U \quad (2.16)$$

In this case,

$$\|H_{\Delta A}(s)\|_\infty \leq \gamma, \quad (\Delta A, \Delta B, \Delta C) \in U, \quad (2.17)$$

and

$$\bar{Q}_{\Delta A} \leq Q, \quad (\Delta A, \Delta B, \Delta C) \in U, \quad (2.18)$$

where $\bar{Q}_{\Delta A}$ is given by (2.12). Consequently,

$$J(A_c, B_c, C_c) \leq J(A_c, B_c, C_c, Q), \quad (2.19)$$

where

$$J(A_c, B_c, C_c, Q) \triangleq \text{tr } Q \bar{R}. \quad (2.20)$$

Proof. First note for clarity that in (2.13) Q denotes an arbitrary element of $\mathbb{N}^{\bar{n}}$ since (2.13) holds for all $Q \in \mathbb{N}^{\bar{n}}$, while in (2.14) Q denotes a specific

solution to (2.14). Now for $(\Delta A, \Delta B, \Delta C) \in U$, (2.14) is equivalent to

$$0 = (\bar{A} + \Delta \bar{A}) \bar{Q} + \bar{Q} (\bar{A} + \Delta \bar{A})^T + \gamma^{-2} \bar{Q} \bar{R} \bar{Q} \\ + \Omega(B_c, C_c, Q) - (\Delta \bar{A} \bar{Q} + \bar{Q} \Delta \bar{A}^T) + \bar{V}. \quad (2.21)$$

Hence, by assumption, (2.21) has a solution $\bar{Q} \in \mathbb{N}^{\bar{n}}$ for all $(\Delta A, \Delta B, \Delta C) \in U$ and, by (2.13), $\Omega(B_c, C_c, Q) - (\Delta \bar{A} \bar{Q} + \bar{Q} \Delta \bar{A}^T)$ is nonnegative definite. Now if the stabilizability condition (2.15) holds for all $(\Delta A, \Delta B, \Delta C) \in U$, it follows from Lemma 12.2 of [5] that $\bar{A} + \Delta \bar{A}$ is asymptotically stable for all $(\Delta A, \Delta B, \Delta C) \in U$.

Conversely, if $\bar{A} + \Delta \bar{A}$ is asymptotically stable for all $(\Delta A, \Delta B, \Delta C) \in U$, then (2.15) holds. The proof of (2.17) now follows from a standard manipulation of (2.14). Next, subtracting (2.12) from (2.20) yields

$$0 = (\bar{A} + \Delta \bar{A})(\bar{Q} - \bar{Q}_{\Delta A}) + (\bar{Q} - \bar{Q}_{\Delta A})(\bar{A} + \Delta \bar{A})^T + \gamma^{-2} \bar{Q} \bar{R} \bar{Q} \\ + \Omega(B_c, C_c, Q) - (\Delta \bar{A} \bar{Q} + \bar{Q} \Delta \bar{A}^T),$$

or, equivalently, since $\bar{A} + \Delta \bar{A}$ is asymptotically stable for all $(\Delta A, \Delta B, \Delta C) \in U$,

$$Q - \bar{Q}_{\Delta A} = \int_0^\infty e^{(\bar{A} + \Delta \bar{A})t} [\gamma^{-2} \bar{Q} \bar{R} \bar{Q} + \Omega(B_c, C_c, Q) \\ - (\Delta \bar{A} \bar{Q} + \bar{Q} \Delta \bar{A}^T)] e^{(\bar{A} + \Delta \bar{A})^T t} dt \geq 0$$

which implies (2.18). The performance bound (2.19) is now an immediate consequence of (2.18). \square

Remark 2.2. Note that (2.15) is actually a closed-loop "disturbability" condition which is not concerned with control as such. This condition guarantees that the closed-loop system does not possess unstable undisturbed modes. In applying Lemma 2.1 it may be convenient to replace condition (2.15) with a stronger condition which is easier to verify in practice. Clearly, (2.15) is satisfied if $[\bar{V} + \gamma^{-2} \bar{Q} \bar{R} \bar{Q} + \Omega(B_c, C_c, Q) - (\Delta \bar{A} \bar{Q} + \bar{Q} \Delta \bar{A}^T)]$ is positive definite for all $(\Delta A, \Delta B, \Delta C) \in U$. This will be the case, for example, if either \bar{V} is positive definite or strict inequality holds in (2.13). Also, it follows from Theorem 3.6 of [5] that (2.15) is implied by the stronger condition that

$$(\bar{A} + \Delta \bar{A}, \bar{D}) \text{ is stabilizable, } (\Delta A, \Delta B, \Delta C) \in U. \quad (2.22)$$

III. The Auxiliary Minimization Problem

As shown in the previous section, the replacement of (2.12) by (2.14) enforces the H_∞ disturbance attenuation constraint and yields an upper bound for the worst case L_2 performance criterion. That is, given a compensator (A_c, B_c, C_c) for which there exists a nonnegative-definite solution to (2.14), the actual worst case L_2 performance (A_c, B_c, C_c) of the compensator is guaranteed to be no worse than the bound given by $J(A_c, B_c, C_c, Q)$. Hence, $J(A_c, B_c, C_c, Q)$ can be interpreted as an auxiliary cost which leads to the following optimization problem.

Auxiliary Minimization Problem. Determine (A_c, B_c, C_c, Q) which minimizes $J(A_c, B_c, C_c, Q)$ subject to (2.14) with $Q \in \mathbb{N}^{\bar{n}}$.

It follows from Lemma 2.1 that the satisfaction of (2.14) for $Q \in \mathbb{N}^{\bar{n}}$ along with the generic condition

(2.15) leads to 1) closed-loop stability for all $(\Delta A, \Delta B, \Delta C) \in U$; 2) prespecified H_∞ performance attenuation for all $(\Delta A, \Delta B, \Delta C) \in U$; and 3) an upper bound for the worst case L_2 performance criterion. Hence, it remains to determine (A_c, B_c, C_c) which minimizes $J(A_c, B_c, C_c, Q)$ and thus provides an optimized bound for the actual worst case L_2 performance $J(A_c, B_c, C_c)$ over all $(\Delta A, \Delta B, \Delta C) \in U$.

IV. Uncertainty Structure: Linear Bound

Having established the theoretical basis for our approach, we now assign explicit structure to the set U and bounding function Ω . Specifically, the uncertainty set U is assumed to be of the form

$$U = \{(\Delta A, \Delta B, \Delta C): \Delta A = \sum_{i=1}^p \sigma_i A_i, \Delta B = \sum_{i=1}^p \sigma_i B_i, \Delta C = \sum_{i=1}^p \sigma_i C_i, \sum_{i=1}^p \sigma_i^2 / a_i^2 \leq 1\}, \quad (4.1)$$

where, for $i = 1, \dots, p$: $A_i \in \mathbb{R}^{n \times n}$, $B_i \in \mathbb{R}^{n \times m}$, and $C_i \in \mathbb{R}^{l \times n}$ are fixed matrices denoting the structure of the parametric uncertainty; a_i is a given positive number; and σ_i is an uncertain real parameter. Note that the uncertain parameters σ_i are assumed to lie in a specified ellipsoidal region in \mathbb{R}^p . The closed-loop system (2.8) thus has structured uncertainty of the form

$$\Delta \bar{A} = \sum_{i=1}^p \sigma_i \bar{A}_i, \quad (4.2)$$

where

$$\bar{A}_i \triangleq \begin{bmatrix} A_i & B_i C_c \\ B_c C_i & 0 \end{bmatrix}; \quad i = 1, \dots, p.$$

Note that the symmetry of the uncertainty set entails no loss of generality by requiring only a redefinition of the nominal plant matrices.

In order to obtain explicit gain expressions for (A_c, B_c, C_c) in Sections 5 and 6, we shall require that at most one of the perturbations ΔB and ΔC is nonzero. We thus consider the cases $(\Delta A, \Delta C) \in U$ or $(\Delta A, \Delta B) \in U$. If this assumption is not imposed, then optimality conditions can still be derived, but at the expense of closed-form gain expressions. In this section, and Section 5 we will assume that $\Delta B = 0$ (i.e., $B_i = 0$, $i = 1, \dots, p$) and $\Omega(B_c, C_c, Q)$ becomes $\Omega(B_c, Q)$. The dual case $\Delta B \neq 0$, $\Delta C = 0$ (i.e., $C_i = 0$, $i = 1, \dots, p$) will be considered in Section 6.

For the structure of U specified by (4.1), the bound Ω satisfying (2.13) can now be given a concrete form.

Proposition 4.1. Let α be an arbitrary positive scalar. Then the function

$$\Omega(B_c, Q) \triangleq \alpha Q + \alpha^{-1} \sum_{i=1}^p a_i^2 \bar{A}_i Q \bar{A}_i^T \quad (4.3)$$

satisfies (2.13) with U given by (4.1) and $\Delta B = 0$.

Proof. See [4]. \square

Remark 4.1. Note that the bound Ω given by (4.3) consists of two distinct terms. The first term αQ can be thought of as arising from an exponential time weighting of the cost, or, equivalently, from a uniform

right shift of the open-loop dynamics. The second term $\alpha^{-1} \sum_{i=1}^p a_i^2 \bar{A}_i Q \bar{A}_i^T$ arises naturally from a multiplicative white noise model. Such interpretations have no bearing on the results obtained here since, only the bound Ω defined by (4.3) is required. We call (4.3) the linear bound since it is linear in Q . For a more detailed discussion on (4.3) see [4].

With Ω defined by (4.3), the modified Riccati equation (2.14) becomes

$$0 = \bar{A}Q + Q\bar{A}^T + \gamma^{-2} Q \bar{R}_\infty Q + \alpha Q + \alpha^{-1} \sum_{i=1}^p a_i^2 \bar{A}_i Q \bar{A}_i^T + \bar{V} \quad (4.4)$$

or, equivalently,

$$0 = \bar{A}_\alpha Q + Q \bar{A}_\alpha^T + \gamma^{-2} Q \bar{R}_\infty Q + \sum_{i=1}^p \delta_i \bar{A}_i Q \bar{A}_i^T + \bar{V}, \quad (4.5)$$

where $\delta_i \triangleq a_i^2 / \alpha$ and

$$\bar{A}_\alpha \triangleq \bar{A} + \frac{\alpha}{2} I_n = \begin{bmatrix} A & B C_c \\ B_c C & A_c \end{bmatrix}.$$

V. Sufficient Conditions for Robust Stability and Performance with Robust H_∞ Disturbance Attenuation: Linear Bound

In this section we state sufficient conditions for characterizing fixed-order (i.e., full- and reduced-order) controllers guaranteeing closed-loop stability for all $(\Delta A, \Delta C) \in U$, constrained H_∞ disturbance attenuation for all $(\Delta A, \Delta C) \in U$, and an optimized worst case L_2 performance bound.

In order to state the main results we require some additional notation and a factorization lemma.

Lemma 5.1. Let $\hat{Q}, \hat{P} \in \mathbb{M}^n$ and suppose $\text{rank } \hat{Q}\hat{P} = n_c$ then there exist $n_c \times n$ G, Γ and $n_c \times n_c$ invertible M , unique except for a change of basis in \mathbb{R}^{n_c} , such that

$$\hat{Q}\hat{P} = G^T M \Gamma, \quad \Gamma G^T = I_{n_c}. \quad (5.1), (5.2)$$

Furthermore, the $n \times n$ matrices

$$\tau \triangleq G^T \Gamma, \quad \tau_\perp \triangleq I_n - \tau \quad (5.3), (5.4)$$

are idempotent and rank n_c and $n - n_c$, respectively. Finally, if $P \in \mathbb{M}^n$ and $\beta \geq 0$ then the inverse

$$S \triangleq (I_n + \beta^2 \gamma^{-2} \hat{Q}P)^{-1} \quad (5.5)$$

exists.

For arbitrary $Q, \hat{Q} \in \mathbb{R}^{n \times n}$ and $\alpha > 0$ define the following notation:

$$\begin{aligned} v_{2s} &\triangleq v_2 + \sum_{i=1}^p \delta_i C_i (Q + \hat{Q}) C_i^T, \\ q_s &\triangleq Q C^T + \sum_{i=1}^p \delta_i A_i (Q + \hat{Q}) C_i^T, \quad \Sigma \triangleq B R^{-1} B^T. \end{aligned}$$

Theorem 5.1. Suppose there exist $Q, P, \hat{Q}, \hat{P} \in \mathbb{N}^n$ satisfying

$$0 = A_\alpha Q + QA_\alpha^T + \gamma^{-2}QR_{1\infty}Q + V_1 + \sum_{i=1}^p \delta_i A_i (Q + \hat{Q})A_i^T - Q_s V_{2s}^{-1} Q_s^T + \tau_\perp Q_s V_{2s}^{-1} Q_s^T \tau_\perp^T, \quad (5.6)$$

$$0 = (A_\alpha + \gamma^{-2}[Q + \hat{Q}]R_{1\infty})^T P + P(A_\alpha + \gamma^{-2}[Q + \hat{Q}]R_{1\infty}) + R_1 + \sum_{i=1}^p \delta_i [A_i^T P A_i + (A_i - Q_s V_{2s}^{-1} C_i)^T \hat{P} (A_i - Q_s V_{2s}^{-1} C_i)] - S^T P E P S + \tau_\perp^T S^T P E P S \tau_\perp, \quad (5.7)$$

$$0 = (A_\alpha - \Sigma P S + \gamma^{-2}QR_{1\infty})\hat{Q} + \hat{Q}(A_\alpha - \Sigma P S + \gamma^{-2}QR_{1\infty})^T + \gamma^{-2}\hat{Q}(R_{1\infty} + \beta^2 S^T P E P S)\hat{Q} + Q_s V_{2s}^{-1} Q_s^T - \tau_\perp Q_s V_{2s}^{-1} Q_s^T \tau_\perp^T, \quad (5.8)$$

$$0 = (A_\alpha - Q_s V_{2s}^{-1} C + \gamma^{-2}QR_{1\infty})^T \hat{P} + \hat{P}(A_\alpha - Q_s V_{2s}^{-1} C + \gamma^{-2}QR_{1\infty}) + S^T P E P S - \tau_\perp^T S^T P E P S \tau_\perp, \quad (5.9)$$

$$\text{rank } \hat{Q} = \text{rank } \hat{P} = \text{rank } \hat{Q} \hat{P} = n_c, \quad (5.10)$$

and let (A_c, B_c, C_c, Q) be given by

$$A_c = \Gamma(A - \Sigma P S - Q_s V_{2s}^{-1} C + \gamma^{-2}QR_{1\infty})G^T \quad (5.11)$$

$$B_c = \Gamma Q_s V_{2s}^{-1}, \quad (5.12)$$

$$C_c = -R_2^{-1} B^T P S G^T, \quad (5.13)$$

$$Q = \begin{bmatrix} Q + \hat{Q} & \hat{Q} \Gamma^T \\ \Gamma Q & \Gamma \hat{Q} \Gamma^T \end{bmatrix}.$$

Then, $(\bar{A} + \Delta \bar{A}, \bar{D})$ is stabilizable if and only if $\bar{A} + \Delta \bar{A}$ is asymptotically stable for all $(\Delta A, \Delta C) \in U$. In this case, the closed-loop transfer function $H_{\Delta \bar{A}}(s)$ satisfies the H_∞ disturbance attenuation constraint

$$\|H_{\Delta \bar{A}}(s)\|_\infty \leq \gamma, \quad (\Delta A, \Delta C) \in U, \quad (5.15)$$

and the worst case L_2 performance criterion (2.10) satisfies the bound

$$J(A_c, B_c, C_c) \leq \text{tr}[(Q + \hat{Q})R_1 + \hat{Q} S^T P E P S \hat{Q}]. \quad (5.16)$$

Proof. The proof follows from Lemma 2.1 by combining the proofs of Theorem 6.1 of [1] and Theorem 6.1 of [4]. \square

Remark 5.1. Theorem 5.1 presents sufficient conditions for designing controllers yielding robust stability and performance with a constraint on the H_∞ norm of the closed-loop transfer function for a state-space system with structured real-valued plant parameter variations. These sufficient conditions comprise a system of three modified Riccati equations and one modified Lyapunov equation coupled by the optimal projection τ , the linear uncertainty bound, and the H_∞ constraint. If the uncertainty bound is deleted, then the results of [1] are recovered. If, alternatively, the uncertainty terms are retained but the H_∞ constraint is sufficiently relaxed, i.e., $\gamma \rightarrow \infty$, then the results of [4] are recovered for the case $B_i = 0$, $i = 1, \dots, p$.

Remark 5.2. To specialize Theorem 5.1 to the full order case $n_c = n$, it is only necessary to set $G^T = \Gamma^{-1}$ so that $G = \Gamma = \tau = I_n$ and $\tau_\perp = 0$ without loss of generality. Now the last term in each of (5.6) - (5.9) can be deleted and G and Γ in (5.11) - (5.14) can be taken to be the identity. It is interesting to note that in the full-order case the H_∞ design problem with structured parameter variations is comprised of four coupled Riccati/Lyapunov equations. This coupling illustrates the breakdown of regulator/estimator separation and shows that the certainty equivalence principle is no longer valid. This is not surprising since separation also breaks down for the full-order L_2 result with parameter uncertainties [4].

Remark 5.3. When solving (5.6) - (5.10) numerically, the uncertainty terms and the H_∞ constraint can be adjusted to examine tradeoffs among performance, robustness, and disturbance rejection. Specifically, the uncertainty range a_i and the structure matrices A_i, C_i appearing in Q_s and V_{2s} along with γ can be varied systematically to determine the region of solvability of the design equations (5.6) - (5.9).

Remark 5.4. Although equations (5.6) - (5.10) appear formidable, they are, in fact, quite numerically tractable. For related problems involving coupled Riccati equations, homotopic continuation methods have been shown to be effective (see [1] and the References therein).

Remark 5.5. We point out that if $\beta = 0$ or, equivalently, $E_{2\infty} = 0$, which corresponds to the "cheap" H_∞ control case (i.e., H_∞ attenuation between disturbances and controls is not constrained), it is possible to obtain closed-form gains (A_c, B_c, C_c) given by a modified set of design equations when all three of $\Delta A, \Delta B$, and ΔC are nonzero. Because of space limitations this result will be given in a future paper.

Remark 5.6. An important special case of the results of Section 5 is obtained by setting $\Delta A = 0$, $\Delta B = 0$, $\Delta C = \sigma_1 C_1$, and $C_1 = C$. The resulting L_2/H_∞ design possesses guaranteed gain margin of $\pm 100\sigma_1$ percent at the sensor output.

VI. The Dual Case: Linear Bound

Unlike the standard LQG result involving a pair of uncoupled Riccati equations, the new result guaranteeing robust stability, robust performance, and disturbance rejection involves a coupled system of four modified Riccati/Lyapunov equations. The asymmetry of these equations suggests the possibility of a dual result in which the modifications to the standard Riccati equations are reversed. One motivation for developing such dual results is that for certain problems the dual bounds may be sharper than the primal bound introduced in Section 4. Furthermore, the dual theory permits distinct H_∞ disturbance weights ($V_{1\infty}$ and $V_{2\infty}$), although we now require $R_{1\infty} = R_2$. Finally, the dual theory allows for uncertainty in the control matrix B (i.e., $\Delta B \neq 0$). Although we now require $\Delta C = 0$, (i.e., $C_i = 0$, $i = 1, \dots, p$) to obtain closed-form gain expressions for (A_c, B_c, C_c) . We begin with the following lemma:

Lemma 6.1. Suppose the system (2.8) is asymptotically stable for all $(\Delta A, \Delta B, \Delta C) \in U$ for a given (A_c, B_c, C_c) . Then

$$J(A_c, B_c, C_c) = \sup_{(\Delta A, \Delta B, \Delta C) \in U} \text{tr } \bar{P}_{\Delta \bar{A}} \bar{V}. \quad (6.1)$$

where $\bar{P}_{\Delta\bar{A}} \in \mathbb{N}^n$ is the unique solution to

$$0 = (\bar{A} + \Delta\bar{A})^T \bar{P}_{\Delta\bar{A}} + \bar{P}_{\Delta\bar{A}} (\bar{A} + \Delta\bar{A}) + \bar{R}. \quad (6.2)$$

Proof. See [3]. \square

Utilizing (6.1) in place of (2.10), the H_∞ disturbance attenuation constraint from plant and sensor disturbances to the state and control variables given by

$$\|\bar{H}_{\Delta\bar{A}}(s)\|_\infty = \|\bar{E}[sI_n - (\bar{A} + \Delta\bar{A})]^{-1}\bar{D}_\infty\|_\infty \leq \gamma \quad (6.3)$$

can now be enforced by replacing (2.14) by the modified Riccati equation

$$0 = \bar{A}^T P + P\bar{A} + \gamma^{-2} P \bar{V}_\infty P + \hat{\Omega}(C_c, P) + \bar{R}, \quad (\Delta A, \Delta B, \Delta C) \in U, \quad (6.4)$$

where

$$\Delta\bar{A}^T P + P\Delta\bar{A} \leq \hat{\Omega}(C_c, P), \quad (\Delta A, \Delta B, \Delta C) \in U. \quad (6.5)$$

Note that (6.4) is merely the dual of (2.14). We also require the condition dual to (2.15) given by

$$([\bar{R} + \gamma^{-2} P \bar{V}_\infty P + \hat{\Omega}(C_c, P) - (\Delta\bar{A}^T P + P\Delta\bar{A})]^{-1/2}, \bar{A} + \Delta\bar{A})$$

is detectable for all $(\Delta A, \Delta B) \in U$. (6.6)

For the structure of U as specified by (4.1) with $\Delta C = 0$, the bound $\hat{\Omega}$ satisfying (6.5) can now be given a concrete form.

Proposition 6.1. Let α be an arbitrary positive scalar. Then the function

$$\hat{\Omega}(C_c, P) \triangleq \alpha P + \alpha^{-1} \sum_{i=1}^p \alpha_i^2 \bar{A}_i^T P \bar{A}_i \quad (6.7)$$

satisfies (6.5) with U given by (4.1) and $\Delta C = 0$. with $\hat{\Omega}$ defined by (6.7), the modified dual Riccati equation (6.4) becomes

$$0 = \bar{A}^T P + P\bar{A} + \gamma^{-2} P \bar{V}_\infty P + \sum_{i=1}^p \delta_i \bar{A}_i^T P \bar{A}_i + \bar{R}. \quad (6.8)$$

We can now state sufficient conditions for robust stability, robust L_2 performance, and robust disturbance attenuation for the dual problem. For arbitrary $Q, P, \hat{P} \in \mathbb{R}^{n \times n}$ and $\alpha > 0$ define the following notation:

$$R_{2s} \triangleq R_2 + \sum_{i=1}^p \delta_i B_i^T (P + \hat{P}) B_i, \quad P_s \triangleq B^T P + \sum_{i=1}^p \delta_i B_i^T (P + \hat{P}) A_i, \\ \hat{S} \triangleq (I_n + \gamma^{-2} \beta^2 Q \hat{P})^{-1}, \quad \bar{S} \triangleq C^T V_2^{-1} C.$$

Theorem 6.1. Suppose there exist $P, Q, \hat{P}, \hat{Q} \in \mathbb{N}^n$ satisfying (5.10) and

$$0 = \bar{A}^T P + P\bar{A} + \gamma^{-2} P \bar{V}_\infty P + R_1 + \sum_{i=1}^p \delta_i \bar{A}_i^T (P + \hat{P}) \bar{A}_i \\ - P_s^T R_{2s}^{-1} P_s + \tau_1^T P_s^T R_{2s}^{-1} P_s \tau_1, \quad (6.9)$$

$$0 = (A_\alpha + \gamma^{-2} V_{1\infty} [P + \hat{P}]) Q + Q (A_\alpha + \gamma^{-2} V_{1\infty} [P + \hat{P}])^T \\ + V_1 + \sum_{i=1}^p \delta_i [A_i Q A_i^T + (A_i - B_i R_{2s}^{-1} P_s) \hat{Q} (A_i - B_i R_{2s}^{-1} P_s)^T] \\ - \hat{S} Q \bar{S} Q \bar{S}^T + \tau_1 \hat{S} Q \bar{S} Q \bar{S}^T \tau_1^T, \quad (6.10)$$

$$0 = (A_\alpha - \hat{S} Q \bar{S} + \gamma^{-2} V_{1\infty} P)^T \hat{P} + \hat{P} (A_\alpha - \hat{S} Q \bar{S} + \gamma^{-2} V_{1\infty} P) + \\ \gamma^{-2} \hat{P} (V_{1\infty} + \beta^2 \hat{S} Q \bar{S} Q \bar{S}^T) \hat{P} + P_s^T R_{2s}^{-1} P_s - \tau_1^T P_s^T R_{2s}^{-1} P_s \tau_1, \quad (6.11)$$

$$0 = (A_\alpha - B R_{2s}^{-1} P_s + \gamma^{-2} V_{1\infty} P) \hat{Q} + \hat{Q} (A_\alpha - B R_{2s}^{-1} P_s + \gamma^{-2} V_{1\infty} P)^T \\ + \hat{S} Q \bar{S} Q \bar{S}^T - \tau_1^T \hat{S} Q \bar{S} Q \bar{S}^T \tau_1^T, \quad (6.12)$$

and let (A_c, B_c, C_c, P) be given by

$$A_c = \Gamma(A - \hat{S} Q \bar{S} - B R_{2s}^{-1} P_s + \gamma^{-2} V_{1\infty} P) G^T, \quad (6.13)$$

$$B_c = \Gamma \hat{S} Q C^T V_2^{-1}, \quad (6.14)$$

$$C_c = -R_{2s}^{-1} P_s G^T, \quad (6.15)$$

$$P = \begin{bmatrix} P + \hat{P} & -\hat{P} G^T \\ -G \hat{P} & G \hat{P} G^T \end{bmatrix}. \quad (6.16)$$

Then, $(\bar{E}, \bar{A} + \Delta\bar{A})$ is detectable if and only if $\bar{A} + \Delta\bar{A}$ is asymptotically stable for all $(\Delta A, \Delta B) \in U$. In this case, the closed-loop transfer function $\bar{H}_{\Delta\bar{A}}(s)$ satisfies the H_∞ disturbance attenuation constraint,

$$\|\bar{H}_{\Delta\bar{A}}(s)\|_\infty \leq \gamma, \quad (\Delta A, \Delta B) \in U, \quad (6.17)$$

and the worst case L_2 performance criterion (6.1) satisfies the bound

$$(A_c, B_c, C_c) \leq \text{tr}[(P + \hat{P}) V_1 + \hat{P} \hat{S} Q \bar{S} Q \bar{S}^T \hat{P}]. \quad (6.18)$$

Remark 6.1. The dual case of Remark 5.6 is obtained by setting $\Delta A = 0$, $\Delta B = \sigma_1 B_1$, $\Delta C = 0$, and $B_1 = B$. The resulting L_2/H_∞ design possesses guaranteed gain margin of $\pm 100\sigma_1$ percent at the input.

VII. Uncertainty Structure and Sufficient Conditions for Robust Stability and Performance with H_∞ Disturbance Attenuation: Quadratic Bound

We now assign a different structure to the uncertainty set U and the bounding function Ω . Specifically, the uncertainty set U is assumed to be of the form

$$U = \{(\Delta A, \Delta B, \Delta C): \Delta A = \sum_{i=1}^p F_i M_i N_i G_i, \\ \Delta B = \sum_{i=1}^p F_i M_i N_i H_i, \Delta C = \sum_{i=1}^p K_i M_i N_i G_i, \\ M_i M_i^T \leq \bar{M}_i, N_i^T N_i \leq \bar{N}_i, i = 1, \dots, p\}, \quad (7.1)$$

where, for $i = 1, \dots, p$: $F_i \in \mathbb{R}^{n \times r_i}$, $G_i \in \mathbb{R}^{l \times n}$, $H_i \in \mathbb{R}^{l \times m}$, and $K_i \in \mathbb{R}^{l \times r_i}$ are fixed matrices denoting the structure of the uncertainty: $\bar{M}_i \in \mathbb{N}^{r_i}$ and $\bar{N}_i \in \mathbb{N}^{l_i}$

are given uncertainty bounds; and $M_i \in \mathbb{R}^{r \times s_i}$ and $N_i \in \mathbb{R}^{s_i \times t_i}$ are uncertain matrices.

In order to obtain explicit gain expressions (A_c , B_c , C_c) we again consider two cases, 1) $(\Delta A, \Delta C) \in U$ with $\Delta B = 0$ and 2) $(\Delta A, \Delta B) \in U$ with $\Delta C = 0$. When $\Delta B = 0$ the closed-loop system has structured uncertainty of the form

$$\Delta \bar{A} = \sum_{i=1}^p \bar{F}_i M_i N_i \bar{G}_i, \quad (7.2)$$

where,

$$\bar{F}_i \triangleq \begin{bmatrix} F_i \\ B_c K_i \end{bmatrix}, \quad \bar{G}_i \triangleq [G_i \ 0].$$

In this case the quadratic bound Ω satisfying (2.13) can now be given a concrete form.

Proposition 7.1. The function

$$\Omega(B_c, Q) \triangleq \sum_{i=1}^p \bar{F}_i \bar{M}_i \bar{F}_i^T + Q \bar{G}_i^T \bar{N}_i \bar{G}_i Q \quad (7.3)$$

satisfies (2.13) with U given by (7.1) and $\Delta B = 0$.

Proof. See [3]. \square

Thus, with Ω defined by (7.3), the modified Riccati equation (7.3) becomes

$$0 = \bar{A}Q + Q\bar{A}^T + \gamma^{-2}Q\bar{R}_\infty Q + \bar{V} + \sum_{i=1}^p [\bar{F}_i \bar{M}_i \bar{F}_i^T + Q \bar{G}_i^T \bar{N}_i \bar{G}_i Q]. \quad (7.4)$$

For arbitrary $Q \in \mathbb{R}^{n \times n}$ define:

$$\begin{aligned} Q_a &\triangleq Q C^T + \sum_{i=1}^p F_i \bar{M}_i K_i, & D &\triangleq \sum_{i=1}^p F_i \bar{M}_i F_i^T, \\ V_{2a} &\triangleq V_2 + \sum_{i=1}^p K_i \bar{M}_i K_i^T, & E &\triangleq \sum_{i=1}^p G_i^T \bar{N}_i G_i. \end{aligned}$$

Theorem 7.1. Suppose these exist Q , P , \hat{Q} , $\hat{P} \in \mathbb{N}^n$ satisfying (5.10) and

$$0 = A Q + Q A^T + \gamma^{-2} Q R_{1\infty} Q + V_1 + Q E Q + D - Q_a V_{2a}^{-1} Q_a^T + \tau_1 Q_a V_{2a}^{-1} Q_a^T \tau_1^T, \quad (7.5)$$

$$\begin{aligned} 0 &= (A + [Q + \hat{Q}][\gamma^{-2} R_{1\infty} + E])^T P \\ &\quad + P(A + [Q + \hat{Q}][\gamma^{-2} R_{1\infty} + E]) \\ &\quad + R_1 - S^T P \Sigma P S + \tau_1^T S^T P \Sigma P S \tau_1, \end{aligned} \quad (7.6)$$

$$\begin{aligned} 0 &= (A - \Sigma P S + Q[\gamma^{-2} R_{1\infty} + E])\hat{Q} + \hat{Q}(A - \Sigma P S \\ &\quad + Q[\gamma^{-2} R_{1\infty} + E])^T + \hat{Q}(\gamma^{-2}[R_{1\infty} + \beta^2 S^T P \Sigma P S] + E)\hat{Q} \\ &\quad + Q_a V_{2a}^{-1} Q_a^T - \tau_1 Q_a V_{2a}^{-1} Q_a^T \tau_1^T, \end{aligned} \quad (7.7)$$

$$\begin{aligned} 0 &= (A - Q_a V_{2a}^{-1} C + Q[\gamma^{-2} R_{1\infty} + E])^T \hat{P} + \hat{P}(A - Q_a V_{2a}^{-1} C \\ &\quad + Q[\gamma^{-2} R_{1\infty} + E]) + S^T P \Sigma P S - \tau_1^T S^T P \Sigma P S \tau_1, \end{aligned} \quad (7.8)$$

and let Q be given by (5.14) and (A_c , B_c , C_c) by

$$A_c = \Gamma(A - \Sigma P S - Q_a V_{2a}^{-1} C + Q[\gamma^{-2} R_{1\infty} + E])G^T, \quad (7.9)$$

$$B_c = \Gamma Q_a V_{2a}^{-1}, \quad (7.10)$$

$$C_c = -R_2^{-1} B^T P S G^T. \quad (7.11)$$

Then, $(\bar{A} + \Delta \bar{A}, \bar{D})$ is stabilizable if and only if $\bar{A} + \Delta \bar{A}$ is asymptotically stable for all $(\Delta A, \Delta C) \in U$. In this case, the closed-loop transfer function $H_{\Delta \bar{A}}(s)$ satisfies the H_∞ disturbance attenuation constraint

$$\|H_{\Delta \bar{A}}(s)\|_\infty \leq \gamma, \quad (\Delta A, \Delta C) \in U, \quad (7.12)$$

and the worst case L_2 performance criterion (2.10) satisfies the bound

$$J(A_c, B_c, C_c) \leq \text{tr}[(Q + \hat{Q})R_1 + \hat{Q}S^T P \Sigma P S \hat{Q}]. \quad (7.13)$$

Proof. The proof follows by combining the proofs of Theorems 6.1 of [1] and Theorem 8.1 of [3]. \square

Remark 7.2. It is interesting to note that the full-order case $n_c = n$ with $G = \Gamma = \tau = I_n$ and $\tau_1 = 0$

(see Remark 5.1), \hat{P} plays no role so that (7.8) is superfluous. Thus, unlike the full-order result for the linear bound involving four equations, the full-order quadratic bound involves three modified Riccati equations coupled by the uncertainty term and the H_∞ constraint. If, alternatively, the reduced-order constraint is retained, but the uncertainty terms are deleted, then the results of [1] are recovered. If, furthermore, the uncertainty terms are retained, but the H_∞ constraint is sufficiently relaxed, i.e., $\gamma \rightarrow \infty$, the results of [3] are recovered.

VIII. The Dual Case: Quadratic Bound

For the structure of U as specified by (7.1) with $\Delta C = 0$, the closed-loop system has structured uncertainty of the form

$$\Delta \bar{A} = \sum_{i=1}^p \bar{F}_i M_i N_i \bar{G}_i, \quad (8.1)$$

where

$$\bar{F}_i \triangleq \begin{bmatrix} F_i \\ 0 \end{bmatrix}, \quad \bar{G}_i \triangleq [G_i \ H_i C_c].$$

Proposition 8.1. The function

$$\hat{\Omega}(C_c, P) \triangleq \sum_{i=1}^p \bar{G}_i^T \bar{N}_i \bar{G}_i + P \bar{F}_i \bar{M}_i \bar{F}_i^T P \quad (8.2)$$

satisfies (6.5) with U given by (7.1) and $\Delta C = 0$.

With $\hat{\Omega}$ defined by (8.2), the modified dual equation (6.4) becomes

$$\begin{aligned} 0 &= \bar{A}^T P + P \bar{A} + \gamma^{-2} P \bar{V}_\infty P + \bar{R} \\ &\quad + \sum_{i=1}^p [\bar{G}_i^T \bar{N}_i \bar{G}_i + P \bar{F}_i \bar{M}_i \bar{F}_i^T P]. \end{aligned} \quad (8.3)$$

For arbitrary $P \in \mathbb{R}^{n \times n}$ define:

$$P_a \triangleq B^T P + \sum_{i=1}^p H_i^T \bar{N}_i G_i, \quad R_{2a} \triangleq R_2 + \sum_{i=1}^p H_i^T \bar{N}_i H_i.$$

Theorem 8.1. Suppose there exists $P, Q, \hat{P}, \hat{Q} \in \mathbb{N}$ satisfying (5.10) and

$$0 = A^T P + PA + \gamma^{-2} P V_{1\infty} P + R_1 + E + PDP - P_a^T R_{2a}^{-1} P_a + \tau_1^T P_a^T R_{2a}^{-1} P_a \tau_1^T, \quad (8.4)$$

$$0 = (A + [\gamma^{-2} V_{1\infty} + D][P + \hat{P}])Q + Q(A + [\gamma^{-2} V_{1\infty} + D][P + \hat{P}])^T + V_1 - \hat{S}Q\hat{S}Q\hat{S}^T + \tau_1 \hat{S}Q\hat{S}Q\hat{S}^T \tau_1^T, \quad (8.5)$$

$$0 = (A - \hat{S}Q\hat{S} + [\gamma^{-2} V_{1\infty} + D]P)^T \hat{P} + \hat{P}(A - \hat{S}Q\hat{S} + [\gamma^{-2} V_{1\infty} + D]P) + \hat{P}(\gamma^{-2} [V_{1\infty} + \beta^2 \hat{S}Q\hat{S}Q\hat{S}^T] + D)\hat{P} + P_a^T R_{2a}^{-1} P_a - \tau_1^T P_a^T R_{2a}^{-1} P_a \tau_1^T, \quad (8.6)$$

$$0 = (A - BR_{2a}^{-1} P_a + [\gamma^{-2} V_{1\infty} + D]P)\hat{Q} + \hat{Q}(A - BR_{2a}^{-1} P_a + [\gamma^{-2} V_{1\infty} + D]P) + \hat{S}Q\hat{S}Q\hat{S}^T - \tau_1 \hat{S}Q\hat{S}Q\hat{S}^T \tau_1^T, \quad (8.7)$$

and let P be given by (6.16) and (A_c, B_c, C_c) by

$$A_c = \Gamma(A - \hat{S}Q\hat{S} - BR_{2a}^{-1} P_a + [\gamma^{-2} V_{1\infty} + D]P)G^T, \quad (8.8)$$

$$B_c = \Gamma \hat{S}Q\hat{S}^T V_{21}^{-1}, \quad (8.9)$$

$$C_c = -R_{2a}^{-1} P_a G^T. \quad (8.10)$$

Then $(\bar{E}, \bar{A} + \Delta \bar{A})$ is detectable if and only if $\bar{A} + \Delta \bar{A}$ is asymptotically stable for all $(\Delta A, \Delta B) \in \mathbb{U}$. In this case, the closed-loop transfer function $\hat{H}_{\Delta \bar{A}}(s)$ satisfies the H_∞ disturbance attenuation constraint

$$\|\hat{H}_{\Delta \bar{A}}(s)\|_\infty \leq \gamma, \quad (\Delta A, \Delta B) \in \mathbb{U}, \quad (8.11)$$

and the worst case L_2 performance criterion (6.1) satisfies the bound

$$J(A_c, B_c, C_c) \leq \text{tr}[(P + \hat{P})V_1 + \hat{P}\hat{S}Q\hat{S}Q\hat{S}^T \hat{P}]. \quad (8.12)$$

VIII. Further Extensions

The results of this paper can be readily extended in several directions:

- 1) Mixed bounds, i.e., letting $\Delta A = \Delta A_1 + \Delta A_2$ and bounding ΔA_1 with the linear bound and ΔA_2 with the quadratic bound (this would unify the linear and quadratic bound results)
- 2) L_2 and H_∞ cross weighting terms (e.g., $x^T R_{12} u$) as well as correlated plant disturbance and sensor noise

- 3) nonstrictly proper plant model, i.e., (2.2) replaced by

$$y(t) = (C + \Delta C)x(t) + (D + \Delta D)u(t) + D_2 w(t) \quad (2.2)'$$

- 4) nonstrictly proper controller, i.e., (2.4) replaced by

$$u(t) = C_c x_c(t) + D_c y(t) \quad (2.4)'$$

and the related problems of singular control weighting ($R_2 \geq 0$) and singular measurement noise ($V_2 \geq 0$)

- 5) discrete-time and sampled data design.

References

- [1] D.S. Bernstein and W.M. Haddad, "LQG Control with An H_∞ Performance Bound: A Riccati Equation Approach", *Proc. Amer. Contr. Conf.*, pp. 796-802, Atlanta, GA, June 1988. (to appear in *IEEE Trans. Autom. Contr.*)
- [2] J.C. Doyle, K. Glover, P.P. Khargonekar, and B.A. Francis, "State-Space Solutions to Standard H_2 and H_∞ Control Problems", *Proc. Amer. Contr. Conf.*, pp. 1691-1696, Atlanta, GA, June 1988.
- [3] D.S. Bernstein and W.M. Haddad, "The Optimal Projection Equations with Petersen-Hollot Bounds: Robust Stability and Performance via Fixed-Order Dynamic Compensation for Systems with Structured Real-Valued Parameter Uncertainty", *IEEE Trans. Autom. Contr.*, Vol. AC-33, pp. 578-582, 1988.
- [4] D.S. Bernstein and W.M. Haddad, "Robust Stability and Performance for Fixed-Order Dynamic Compensation with Guaranteed Cost Bounds", *Proc. Amer. Contr. Conf.*, pp. 2471-2476, Atlanta, GA, June 1988.
- [5] W.M. Wonham, "Linear Multivariable Control: A Geometric Approach", *Springer-Verlag*, New York, 1979.

APPENDIX K: Tracking Control

- 67. D. S. Bernstein and W. M. Haddad, "Optimal Output Feedback for Nonzero Set Point Regulation," IEEE Trans. Autom. Contr., Vol. AC-32, pp. 641-645, 1987.
- 86. W. M. Haddad and D. S. Bernstein, "Optimal Output Feedback for Nonzero Set Point Regulation: The Discrete-Time Case," Int. J. Contr., Vol. 47, pp. 529-536, 1988.
- 103. W. M. Haddad and D. S. Bernstein, "Optimal Nonzero Set Point Regulation via Fixed-Order Dynamic Compensation," IEEE Trans. Autom. Contr., Vol. 33, pp. 848-852, 1988.
- 125. D. S. Bernstein and W. M. Haddad, "Optimal Reduced-Order State Estimation for Unstable Plants," submitted to Int. J. Contr.

Optimal Output Feedback for Nonzero Set Point Regulation

DENNIS S. BERNSTEIN AND WASSIM M. HADDAD

Abstract—Motivated by the results of Artstein and Leizarowitz [2] on steady-state periodic tracking, a continuous-time nonzero set point regulation problem is considered which involves 1) noisy and nonnoisy measurements, 2) weighted and unweighted controls, 3) correlated plant/measurement noise and cross weighting, 4) nonzero-mean disturbances, and 5) state-, control-, and measurement-dependent white noise. It is shown that in the absence of multiplicative disturbances the closed-loop control can be designed independently of the open-loop control. Unlike [2], the results are obtained without using the overtaking criterion.

I. INTRODUCTION

The quadratic performance criterion

$$J \triangleq \int_0^{\infty} x^T(t) Q x(t) + u^T(t) R u(t) dt \quad (1.1)$$

expresses the desire to minimize deviations of the state $x(t)$ of the system

$$\dot{x}(t) = Ax(t) + Bu(t) \quad (1.2)$$

from the regulation point $x = 0$. As is well known [1, pp. 270–276], the nonzero set point criterion

$$J_x = \int_0^{\infty} [x(t) - \bar{x}]^T Q [x(t) - \bar{x}] + u^T(t) R u(t) dt \quad (1.3)$$

Manuscript received July 17, 1986; revised March 4, 1987. This work was supported in part by the Air Force Office of Scientific Research under Contracts F49620-86-C-0002 and F49620-86-C-0038.

D. S. Bernstein is with the Government Aerospace Systems Division, Harris Corporation, Melbourne, FL 32902.

W. M. Haddad is with the Department of Mechanical Engineering, Florida Institute of Technology, Melbourne, FL 32901.

IEEE Log Number 8714805.

presents no additional difficulty as long as $x(t)$ and $u(t)$ are replaced by $x(t) - \bar{x}$ and $u(t) - \bar{u}$, where \bar{u} satisfies

$$0 = A\bar{x} + B\bar{u}. \quad (1.4)$$

Closer inspection, however, reveals that this approach is suboptimal. Specifically, the offset \bar{u} in the control may correspond to an unacceptably high level of control effort when $\bar{u}^T R \bar{u}$ is large. Hence, this approach overlooks design tradeoffs concerning the control effort required for maintaining the nonzero regulation point \bar{x} . Moreover, such an approach is impossible when \bar{u} satisfying (1.4) does not exist.

A significant advance in extending the LQR formulation to steady-state tracking problems (and, hence, to nonzero set point regulation) was given by Artstein and Leizarowitz in [2]. They consider the performance criterion

$$J_\infty \triangleq \int_0^\infty [x(t) - \Gamma(t)]^T Q [x(t) - \Gamma(t)] + u^T(t) R u(t) dt \quad (1.5)$$

where $\Gamma(\cdot)$ is periodic on $[0, \infty)$ and the minimization of J_∞ is performed in the sense of the overtaking criterion. For the nonzero set point problem ($\Gamma(t) = \bar{x}$) with full-state feedback plus constant offset control law

$$u(t) = Kx(t) + \alpha \quad (1.6)$$

it follows from [2, Theorem 2] that K and α are given by

$$K = -R^{-1}B^T P, \quad (1.7)$$

$$\alpha = -R^{-1}B^T(A - \Sigma P)^{-1}Q\bar{x} \quad (1.8)$$

where P satisfies the Riccati equation

$$0 = A^T P + PA + Q - P \Sigma P \quad (1.9)$$

with $\Sigma \triangleq BR^{-1}B^T$.

Two features of the control law (1.6)–(1.8) are noteworthy. First, (1.6) consists of both a closed-loop feedback component $Kx(t)$ and an open-loop component α depending upon the regulation point (Fig. 1). Second (and more important), is the observation that the closed-loop control component is independent of the open-loop component. From a practical point of view this feature is quite useful since it implies that the feedback gain K can be determined without regard to the set point. Hence, a change in the desired set point \bar{x} during on-line operation does not necessitate resolving the Riccati equation in real time; only α requires updating. For a new value of \bar{x} , α can readily be recomputed on-line via the matrix multiplication operation (1.8).

The contribution of the present note is an extension of the result of [2] as applied to the nonzero set point regulation problem without using the overtaking criterion. We extend this result in the following different ways.

1) **Output Feedback with Noisy and Nonnoisy Measurements:** To obtain a more realistic problem setting, we consider the case in which the full state is not available, but rather only measured linear combinations of states. Moreover, we consider the possibility that some of the measurements are corrupted by white noise while others are noise free. Note that the noise-free case was considered in [3] while the fully noisy case is the standard assumption in LQG theory. As in [4]–[6] we express the solution in terms of a projection corresponding to the noise-free measurements.

2) **Singular Control Weighting:** As noted in [6], [7] static continuous-time feedback of noise-corrupted measurements results in unbounded cost unless the corresponding controls are unweighted. Hence, we allow for both weighted and unweighted controls to which the noise-free and noisy measurements are fed, respectively. This setting leads to an additional projection dual to the projection arising from the noise-free measurements [6].

3) **Correlated Plant and Measurement Noise and Cross Weighting:** To allow greater design flexibility we allow the possibility that the plant and measurement noise are correlated. In addition, we consider the dual design feature, namely, cross weighting in the performance criterion.

4) **Nonzero-Mean Disturbances:** In addition to the presence of zero-

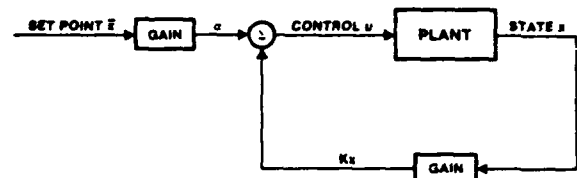


Fig. 1.

mean white plant disturbances we allow for the possibility of a nonzero constant disturbance offset. In contrast to [1, pp. 277–281], our result shows that the presence of a constant disturbance offset leads to an additional offset in the open-loop component of the control.

5) **Multiplicative White Noise:** In addition to the above generalizations we allow for the presence of multiplicative disturbances in the plant. The control law thus generalizes previous results involving state-, control- and measurement-dependent noise [8]–[11]. As shown in [12]–[14], the multiplicative white noise model can be used to guarantee robustness with respect to deterministic plant parameter variations.

II. NOTATION AND DEFINITIONS

| | |
|---|--|
| $\mathbb{R}, \mathbb{R}^{n \times r}, \mathbb{R}^r$ | real numbers, $r \times s$ real matrices, $\mathbb{R}^{r \times 1}$ |
| \mathbb{E} | expectation |
| $I_r, (\cdot)^T$ | $r \times r$ identity, transpose |
| \oplus, \otimes | Kronecker sum, Kronecker product |
| asymptotically stable matrix | matrix with eigenvalues in open left-half plane |
| $n, m_1, m_2, l_1, l_2, p, r$ | positive integers |
| x, u_1, u_2, y_1, y_2 | n, m_1, m_2, l_1, l_2 -dimensional vectors |
| $A, A_i; B_1, B_{1i}; C_1, C_{1i}$ | $n \times n$ matrices; $n \times m_1$ matrices; $l_1 \times n$ matrices, $i = 1, \dots, p$ |
| B_2, C_2, K_1, K_2, L | $n \times m_2, l_2 \times n, m_1 \times l_2, m_2 \times l_1, r \times n$ matrices |
| $\alpha_1, \alpha_2, \gamma, \delta$ | m_1, m_2, n, r -dimensional vectors |
| $w_i(t)$ | unit variance white noise, $i = 1, \dots, p$ |
| $w_0(t), w_1(t)$ | n -dimensional, l_1 -dimensional white noise |
| V_0, V_1 | intensities of w_0, w_1 ; $V_0 \geq 0, V_1 > 0$ |
| V_{01} | $n \times l_1$ cross intensity of w_0, w_1 |
| R_0, R_1 | $r \times r$ and $m_1 \times m_1$ state and control weightings; $R_0 \geq 0, R_1 > 0$ |
| R_{01} | $r \times m_1$ cross weighting; $R_0 - R_{01}R_1^{-1}R_{01}^T \geq 0$ |
| \bar{A}, \bar{A}_i | $A + B_1K_1C_2 + B_2K_2C_1, A_i + B_{1i}K_1C_2 + B_{2i}K_2C_{1i}, i = 1, \dots, p$ |
| \bar{B}, \bar{B}_i | $B_1\alpha_1 + B_2\alpha_2 + \gamma, B_{1i}\alpha_{1i}, i = 1, \dots, p$ |
| $\bar{w}(t)$ | $w_0(t) + B_2K_2w_1(t)$ |
| \bar{V} | $V_0 + V_{01}K_2^T B_2^T + B_2K_2 V_{01}^T + B_2K_2 V_1 K_1^T B_1^T$ |
| \bar{R} | $L^T R_0 L + L^T R_{01} K_1 C_2 + C_2^T K_1^T R_{01}^T L + C_2^T K_1^T R_1 K_1 C_2$ |

For arbitrary $n \times n$ Q, P such that the indicated inverses exist, define:

$$\tau_1 \triangleq QC_2^T(C_2QC_2^T)^{-1}C_2, \quad \tau_2 \triangleq B_2(B_2^TPB_2)^{-1}B_2^TP,$$

$$\tau_{1\perp} \triangleq I_n - \tau_1, \quad \tau_{2\perp} \triangleq I_n - \tau_2,$$

$$V_{1i} \triangleq V_1 + \sum_{j=1}^p C_{1j}(Q + \bar{A}_i\bar{A}_i^T)C_{1j}^T, \quad R_{1i} \triangleq R_1 + \sum_{j=1}^p B_{1j}^TPB_{1j},$$

$$Q_c \triangleq QC_2^T + V_{01} + \sum_{i=1}^p A_i(Q + \bar{A}_i\bar{A}_i^T)C_{1i}^T,$$

$$\Phi_i \triangleq B_1^TP + R_{01}^TL + \sum_{j=1}^p B_{1j}^TPA_{1j},$$

$$\tilde{A}_i \triangleq A - B_i R_{1i}^{-1} \Phi_i \tau_i - \tau_i Q_i V_{1i}^{-1} C_{1i}, \quad \tilde{m}_i \triangleq -\tilde{A}_i^{-1} \tilde{B}_i$$

$$\tilde{A}_{1i} \triangleq A_i - B_{1i} R_{1i}^{-1} \Phi_i \tau_i - \tau_i Q_i V_{1i}^{-1} C_{1i},$$

$$\Phi_i \triangleq R_{1i} - (\Phi_i \tau_i \tilde{A}_i^{-1} B_i + B_i^T \tilde{A}_i^{-T} \tau_i^T \Phi_i^T),$$

$$\Delta_{1i} \triangleq \Phi_i \tau_i \tilde{A}_i^{-1} + B_i^T \tilde{A}_i^{-T} P, \quad \Delta_{2i} \triangleq B_i^T (P \tilde{A}_i^{-1} + \tilde{A}_i^{-T} P),$$

$$\Omega_{1i} \triangleq \Delta_{1i} B_i, \quad \Omega_{2i} \triangleq \Delta_{2i} B_i,$$

$$\Lambda_i \triangleq \tilde{A}_i^{-T} (L^T R_0 - \tau_i^T \Phi_i^T R_{1i}^{-1} R_{0i}^T),$$

$$\Lambda_{1i} \triangleq B_{1i}^T \Lambda_i - R_{0i}^T, \quad \Lambda_{2i} \triangleq B_{2i}^T \Lambda_i,$$

$$Q \triangleq Q C_i^T + V_{0i}, \quad \Phi \triangleq B_i^T P + R_{0i}^T,$$

$$\tilde{A} \triangleq A - B_i R_{1i}^{-1} \Phi_i \tau_i - \tau_i Q_i V_{1i}^{-1} C_{1i}, \quad \tilde{m} \triangleq -\tilde{A}^{-1} \tilde{B},$$

$$\Phi \triangleq R_i - (\Phi_i \tau_i \tilde{A}_i^{-1} B_i + B_i^T \tilde{A}_i^{-T} \tau_i^T \Phi_i^T),$$

$$\Delta_1 \triangleq \Phi_i \tau_i \tilde{A}_i^{-1} + B_i^T \tilde{A}_i^{-T} P, \quad \Delta_2 \triangleq B_i^T (P \tilde{A}_i^{-1} + \tilde{A}_i^{-T} P),$$

$$\Omega_1 \triangleq \Delta_1 B_i, \quad \Omega_2 \triangleq \Delta_2 B_i,$$

$$\Lambda \triangleq \tilde{A}^{-T} (L^T R_0 - \tau_i^T \Phi_i^T R_{1i}^{-1} R_{0i}^T),$$

$$\Lambda_1 \triangleq B_{1i}^T \Lambda - R_{0i}^T, \quad \Lambda_2 \triangleq B_{2i}^T \Lambda,$$

$$\tilde{A}_0 \triangleq A - B_i R_{1i}^{-1} \Phi_i \tau_i, \quad \Phi_0 \triangleq R_i - (\Phi_i \tau_i \tilde{A}_0^{-1} B_i + B_i^T \tilde{A}_0^{-T} \tau_i^T \Phi_i^T),$$

$$\Delta_{10} \triangleq \Phi_i \tau_i \tilde{A}_0^{-1} + B_i^T \tilde{A}_0^{-T} P,$$

$$\Lambda_{10} \triangleq B_{1i}^T \tilde{A}_0^{-T} (L^T R_0 - \tau_i^T \Phi_i^T R_{1i}^{-1} R_{0i}^T) - R_{0i}^T,$$

$$\tilde{A}_1 \triangleq A - B_i R_{1i}^{-1} \Phi_i, \quad \Sigma \triangleq B_i R_{1i}^{-1} B_i^T, \quad A_p \triangleq A - \Sigma P.$$

III. NONZERO SET POINT REGULATION

Nonzero Set Point Problem: Given the controlled system

$$\begin{aligned} \dot{x}(t) = & \left(A + \sum_{i=1}^p v_i(t) A_i \right) x(t) \\ & + \left(B_1 + \sum_{i=1}^p v_i(t) B_{1i} \right) u_1(t) + B_2 u_2(t) + w_0(t) + \gamma \end{aligned} \quad (3.1)$$

with measurements

$$y_1(t) = \left(C_1 + \sum_{i=1}^p v_i(t) C_{1i} \right) x(t) + w_1(t), \quad (3.2)$$

$$y_2(t) = C_2 x(t) \quad (3.3)$$

where $t \in [0, \infty)$, determine K_1, K_2, α_1 , and α_2 such that the static output feedback law

$$u_1(t) = K_1 y_2(t) + \alpha_1, \quad (3.4)$$

$$u_2(t) = K_2 y_1(t) + \alpha_2 \quad (3.5)$$

minimizes the performance criterion

$$\begin{aligned} J(K_1, K_2, \alpha_1, \alpha_2) = & \lim_{t \rightarrow \infty} \mathbb{E} \{ (Lx(t) - \delta)^T R_0 (Lx(t) - \delta) \\ & + 2(Lx(t) - \delta)^T R_{01} u_1(t) + u_1^T(t) R_1 u_1(t) \}. \end{aligned} \quad (3.6)$$

The closed-loop system (3.1)–(3.5) can be written as

$$\dot{x}(t) = \left(\tilde{A} + \sum_{i=1}^p v_i(t) \tilde{A}_i \right) x(t) + \tilde{B} + \sum_{i=1}^p v_i(t) \tilde{B}_i + \tilde{w}(t). \quad (3.7)$$

To analyze (3.7) define the second-moment and covariance matrices

$$\tilde{Q}(t) \triangleq \mathbb{E}[x(t)x^T(t)], \quad Q(t) \triangleq \tilde{Q}(t) - m(t)m^T(t)$$

where $m(t) \triangleq \mathbb{E}[x(t)]$. It follows from [15, p. 142], that $\tilde{Q}(t)$, $Q(t)$, and $m(t)$ satisfy

$$\begin{aligned} \dot{\tilde{Q}}(t) = & \tilde{A} \tilde{Q}(t) + \tilde{Q}(t) \tilde{A}^T + \tilde{V} + \tilde{B} m^T(t) + m(t) \tilde{B}^T \\ & + \sum_{i=1}^p [\tilde{A}_i \tilde{Q}(t) \tilde{A}_i^T + \tilde{A}_i m(t) \tilde{B}_i^T + \tilde{B}_i m^T(t) \tilde{A}_i^T + \tilde{B}_i \tilde{B}_i^T], \end{aligned} \quad (3.8)$$

$$\begin{aligned} \dot{Q}(t) = & \tilde{A} Q(t) + Q(t) \tilde{A}^T + \tilde{V} + \sum_{i=1}^p [\tilde{A}_i Q(t) \tilde{A}_i^T + \tilde{A}_i m(t) m^T(t) \tilde{A}_i^T \\ & + \tilde{A}_i m(t) \tilde{B}_i^T + \tilde{B}_i m^T(t) \tilde{A}_i^T + \tilde{B}_i \tilde{B}_i^T], \end{aligned} \quad (3.9)$$

$$\dot{m}(t) = \tilde{A} m(t) + \tilde{B}. \quad (3.10)$$

To consider the steady state, we restrict our consideration to the set of second-moment stabilizing gains

$$\mathcal{S}_s \triangleq \{(K_1, K_2): \tilde{A} \oplus \tilde{A} + \sum_{i=1}^p \tilde{A}_i \oplus \tilde{A}_i \text{ is asymptotically stable}\}.$$

It follows from fundamental properties of Lyapunov equations that if $(K_1, K_2) \in \mathcal{S}_s$, then \tilde{A} is also asymptotically stable. Hence, for $(K_1, K_2) \in \mathcal{S}_s$, $\tilde{Q} \triangleq \lim_{t \rightarrow \infty} \tilde{Q}(t)$, $Q \triangleq \lim_{t \rightarrow \infty} Q(t)$ and $m \triangleq \lim_{t \rightarrow \infty} m(t)$ exist and satisfy

$$\begin{aligned} 0 = & \tilde{A} \tilde{Q} + \tilde{Q} \tilde{A}^T + \tilde{B} m + m \tilde{B}^T + \tilde{V} + \sum_{i=1}^p [\tilde{A}_i \tilde{Q} \tilde{A}_i^T \\ & + \tilde{A}_i m \tilde{B}_i^T + \tilde{B}_i m^T \tilde{A}_i^T + \tilde{B}_i \tilde{B}_i^T], \end{aligned} \quad (3.11)$$

$$\begin{aligned} 0 = & \tilde{A} Q + Q \tilde{A}^T + \tilde{V} + \sum_{i=1}^p [\tilde{A}_i Q \tilde{A}_i^T + \tilde{A}_i m m^T \tilde{A}_i^T \\ & + \tilde{A}_i m \tilde{B}_i^T + \tilde{B}_i m^T \tilde{A}_i^T + \tilde{B}_i \tilde{B}_i^T], \end{aligned} \quad (3.12)$$

$$0 = \tilde{A} m + \tilde{B}. \quad (3.13)$$

Now $J(K_1, K_2, \alpha_1, \alpha_2)$ is given by

$$\begin{aligned} J(K_1, K_2, \alpha_1, \alpha_2) = & \text{tr} \{ (Q + m m^T) \tilde{R} \} - 2 m^T L^T R_0 \delta \\ & + \delta^T R_0 \delta + 2 m^T L^T R_{01} \alpha_1 \\ & - 2 \delta^T R_{01} K_1 C_2 m - 2 \delta^T R_{01} \alpha_1 \\ & + 2 m^T C_2^T K_1^T R_1 \alpha_1 + \alpha_1^T R_1 \alpha_1. \end{aligned} \quad (3.14)$$

Associated with Q is its dual $P \geq 0$ which is the unique solution of

$$0 = \tilde{A}^T P + P \tilde{A} + \tilde{R} + \sum_{i=1}^p \tilde{A}_i^T P \tilde{A}_i. \quad (3.15)$$

To obtain closed-form expressions for the feedback gains we further restrict consideration to the set

$$\begin{aligned} \mathcal{S}_s^+ \triangleq & \{(K_1, K_2) \in \mathcal{S}_s: C_2 Q C_2^T, B_2^T P B_2, \Phi, \\ & \text{and } \Omega_{1i} \Phi_i^{-1} \Omega_{1i} + \Omega_{2i} \text{ are invertible}\}, \end{aligned}$$

and assume

$$[B_{1i} \neq 0 = C_{1i} = 0], \quad i = 1, \dots, p. \quad (3.16)$$

Optimizing (3.14) subject to (3.12) and (3.13) yields the following result illustrated in Fig. 2.

Theorem 3.1: Suppose $K_1, K_2, \alpha_1, \alpha_2$ solve the nonzero set point

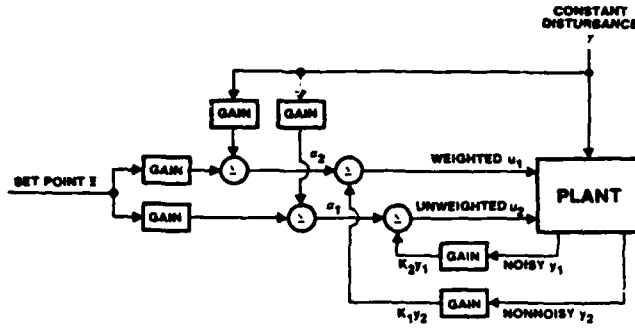


Fig. 2.

problem with $(K_1, K_2) \in \mathcal{S}_*$. Then there exist $n \times n$ $Q, P \geq 0$ such that

$$K_1 = -R_{12}^{-1} \Phi_s Q C_2^T (C_2 Q C_2^T)^{-1}, \quad (3.17)$$

$$K_2 = -(B_2^T P B_2)^{-1} B_2^T P Q V_{1s}^{-1}, \quad (3.18)$$

$$\alpha_1 = \Phi_s^{-1} (\Omega_{1s} \alpha_2 + \Delta_{1s} \gamma - \Lambda_{1s} \delta), \quad (3.19)$$

$$\alpha_2 = (\Omega_{1s}^T \Phi_s^{-1} \Omega_{1s} + \Omega_{2s})^{-1} [(\Omega_{1s}^T \Phi_s^{-1} \Delta_{1s} + \Delta_{2s}) \gamma - (\Omega_{1s}^T \Phi_s^{-1} \Lambda_{1s} + \Lambda_{2s}) \delta], \quad (3.20)$$

and such that Q and P satisfy

$$\begin{aligned} 0 = & (A - B_1 R_{11}^{-1} \Phi_s \tau_1) Q + Q (A - B_1 R_{11}^{-1} \Phi_s \tau_1)^T + V_0 \\ & + \sum_{i=1}^p [(A_i - B_{1i} R_{1i}^{-1} \Phi_s \tau_i) Q (A_i - B_{1i} R_{1i}^{-1} \Phi_s \tau_i)^T \\ & + \hat{A}_{1i} \hat{m}_i \hat{m}_i^T \hat{A}_{1i}^T + \hat{A}_{1i} \hat{m}_i \hat{B}_i^T + \hat{B}_i \hat{m}_i^T \hat{A}_{1i}^T + \hat{B}_i \hat{B}_i^T] \\ & - Q_s V_{1s}^{-1} Q_s^T + \tau_{2s} Q_s V_{1s}^{-1} Q_s^T \tau_{2s}^T, \end{aligned} \quad (3.21)$$

$$\begin{aligned} 0 = & (A - \tau_2 Q_s V_{1s}^{-1} C_1)^T P + P (A - \tau_2 Q_s V_{1s}^{-1} C_1) + R_0 \\ & + \sum_{i=1}^p (A_i - \tau_2 Q_{si} V_{1i}^{-1} C_{1i})^T P (A_i - \tau_2 Q_{si} V_{1i}^{-1} C_{1i}) \\ & - \Phi_s^T R_{1s}^{-1} \Phi_s + \tau_{1s}^T \Phi_s^T R_{1s}^{-1} \Phi_s \tau_{1s}. \end{aligned} \quad (3.22)$$

Outline of Proof: As in [16] the result is obtained by forming the Lagrangian while accounting for (3.12) and (3.13). Define

$$\begin{aligned} \mathcal{L}(K_1, K_2, \alpha_1, \alpha_2, Q, m) = & \text{tr} [\lambda_0 J(K_1, K_2, \alpha_1, \alpha_2) \\ & + P \text{ (RHS of (3.12))} + \lambda^T (\bar{A} m + \bar{B})] \end{aligned}$$

where $\lambda_0 \geq 0$ and $\lambda \in \mathbb{R}^n$. Setting $\partial \mathcal{L} / \partial Q = 0$ and using the second-moment stability assumption it follows that $\lambda_0 = 1$ without loss of generality. The derivation now follows by setting the partial derivatives of \mathcal{L} with respect to $K_1, K_2, \alpha_1, \alpha_2$, and m to zero and solving for the gains. To assist the reader in carrying out the details we note that λ is given by

$$\begin{aligned} \lambda = & -2 \bar{A}^T [-P \bar{B} - \sum_{i=1}^p \bar{A}_i^T P \bar{B}_i + L^T R_0 \delta \\ & - L^T R_{01} \alpha_1 + C_2^T K_1 R_{01}^T \delta - C_2^T K_1^T R_{11} \alpha_1]. \quad \square \end{aligned}$$

Remark 3.1: Because of the presence of δ in (3.21) via α_1 and α_2 in \bar{B} (see the definition of \hat{m}_i) and α_i in \bar{B}_i , the closed-loop component of the control law (3.17)–(3.20) cannot be designed independently of the open-loop component. As now shown, independence is recovered when the multiplicative noise terms are absent.

IV. SPECIALIZATIONS OF THEOREM 3.1

To draw connections with the previous literature, a series of specializations of Theorem 3.1 is now given. We begin by deleting all multiplicative white noise terms, i.e.,

$$A_i, B_{1i}, C_{1i} = 0, \quad i = 1, \dots, p. \quad (4.1)$$

In this case the stabilizing set \mathcal{S}_* can be characterized by

$$\mathcal{S} = \{(K_1, K_2): \bar{A} \text{ is asymptotically stable}\},$$

and, furthermore, \mathcal{S}_* becomes

$$\begin{aligned} \mathcal{S}^+ \triangleq & \{(K_1, K_2) \in \mathcal{S}: C_2 Q C_2^T, B_2^T P B_2, \\ & \Phi \text{ and } \Omega_1^T \Phi^{-1} \Omega_1 + \Omega_2 \text{ are invertible}\}. \end{aligned}$$

Corollary 4.1: Assume (4.1) is satisfied and suppose $K_1, K_2, \alpha_1, \alpha_2$ solve the nonzero set point problem with $(K_1, K_2) \in \mathcal{S}^+$. Then there exist $n \times n$ $Q, P \geq 0$ such that

$$K_1 = -R_1^{-1} \Phi Q C_2^T (C_2 Q C_2^T)^{-1}, \quad (4.2)$$

$$K_2 = -(B_2^T P B_2)^{-1} B_2^T P Q V_1^{-1}, \quad (4.3)$$

$$\alpha_1 = \Phi^{-1} (\Omega_1 \alpha_2 + \Delta_1 \gamma - \Lambda_1 \delta), \quad (4.4)$$

$$\alpha_2 = (\Omega_1^T \Phi^{-1} \Omega_1 + \Omega_2)^{-1} [(\Omega_1^T \Phi^{-1} \Delta_1 + \Delta_2) \gamma - (\Omega_1^T \Phi^{-1} \Lambda_1 + \Lambda_2) \delta], \quad (4.5)$$

and such that Q and P satisfy

$$\begin{aligned} 0 = & (A - B_1 R_1^{-1} \Phi \tau_1) Q + Q (A - B_1 R_1^{-1} \Phi \tau_1)^T + V_0 \\ & - Q V_1^{-1} Q^T + \tau_{2s} Q V_1^{-1} Q^T \tau_{2s}^T, \end{aligned} \quad (4.6)$$

$$\begin{aligned} 0 = & (A - \tau_2 Q V_1^{-1} C_1)^T P + P (A - \tau_2 Q V_1^{-1} C_1) + R_0 \\ & - \Phi^T R_1^{-1} \Phi + \tau_{1s}^T \Phi^T R_1^{-1} \Phi \tau_{1s}. \end{aligned} \quad (4.7)$$

In addition to (4.1) we assume further that all controls are weighted and all measurements are noise free, i.e.,

$$B_2 = 0, \quad C_1 = 0. \quad (4.8)$$

This corresponds to the setting considered in [3]. It follows from the proof of Theorem 3.1 that the assumption $B_2 = 0$ leads to $\tau_2 = 0$, and $C_1 = 0$ corresponds to deleting (3.5). Hence \mathcal{S} and \mathcal{S}^+ are now given by

$$\mathcal{S}_0 \triangleq \{K_1: A + B_1 K_1 C_2 \text{ is asymptotically stable}\},$$

$$\mathcal{S}_0^+ \triangleq \{K_1 \in \mathcal{S}_0: C_2 Q C_2^T \text{ and } \Phi_0 \text{ are invertible}\}.$$

Corollary 4.2: Assume (4.1) and (4.8) are satisfied and suppose K_1 and α_1 solve the nonzero set point problem with $K_1 \in \mathcal{S}_0^+$. Then there exist $n \times n$ $Q, P \geq 0$ such that

$$K_1 = -R_1^{-1} \Phi Q C_2^T (C_2 Q C_2^T)^{-1}, \quad (4.9)$$

$$\alpha_1 = \Phi_0^{-1} (\Delta_{10} \gamma - \Lambda_{10} \delta), \quad (4.10)$$

and such that Q and P satisfy

$$0 = (A - B_1 R_1^{-1} \Phi \tau_1) Q + Q (A - B_1 R_1^{-1} \Phi \tau_1)^T + V_0, \quad (4.11)$$

$$0 = A^T P + P A + R_0 - \Phi^T R_1^{-1} \Phi + \tau_{1s}^T \Phi^T R_1^{-1} \Phi \tau_{1s}. \quad (4.12)$$

We now specialize further to the full-state feedback case, i.e.,

$$C_2 = I_n, \quad (4.13)$$

and hence $\tau_1 = I_n$ and $\tau_{1s} = 0$. Now \mathcal{S}_0 and \mathcal{S}_0^+ become

$$\mathcal{S}_1 \triangleq \{K_1: A + B_1 K_1 \text{ is asymptotically stable}\},$$

$$\mathcal{S}_1^+ \triangleq \{K_1 \in \mathcal{S}_1: Q \text{ is invertible}\}.$$

Corollary 4.3: Assume (4.1), (4.8), and (4.13) are satisfied and suppose K_1 and α_1 solve the nonzero set point problem with $K_1 \in \mathcal{S}_1^+$. Then there exist $n \times n$ Q , $P \geq 0$ such that

$$K_1 = -R_1^{-1}\Phi, \quad (4.14)$$

$$\alpha_1 = R_1^{-1}B_1^T\hat{A}_1^T P \gamma - R_1^{-1}[B_1^T\hat{A}_1^T(L^TR_0 - \Phi^TR_1^{-1}R_0^T) - R_0^T]\delta \quad (4.15)$$

and such that Q and P satisfy

$$0 = (A - B_1R_1^{-1}\Phi)Q + Q(A - B_1R_1^{-1}\Phi)^T + V_0, \quad (4.16)$$

$$0 = A^TP + PA + R_0 - \Phi^TR_1\Phi. \quad (4.17)$$

Finally, setting

$$\gamma = 0, \quad R_0 = 0, \quad L = I_n \quad (4.18)$$

we obtain the result of [2].

Corollary 4.4: Assume (4.1), (4.8), (4.13), and (4.18) are satisfied and suppose K_1 and α_1 solve the nonzero set point problem with $K_1 \in \mathcal{S}_1^+$. Then there exists $n \times n$ $P \geq 0$ such that

$$K_1 = -R_1^{-1}B_1^TP, \quad (4.19)$$

$$\alpha_1 = -R_1^{-1}B_1^TA_1^TR_0\delta \quad (4.20)$$

and such that P satisfies

$$0 = A^TP + PA + R_0 - P\Sigma P. \quad (4.21)$$

REFERENCES

- [1] K. Kwakernaak and R. Sivan, *Linear Optimal Control Systems*. New York: Wiley, 1972.
- [2] Z. Artstein and A. Leizarowitz, "Tracking periodic signals with the overtaking criterion," *IEEE Trans. Automat. Contr.*, vol. AC-30, pp. 1123-1126, 1985.
- [3] W. S. Levine and M. Athans, "On the determination of the optimal constant output feedback gains for linear multivariable systems," *IEEE Trans. Automat. Contr.*, vol. AC-15, pp. 44-48, 1970.
- [4] J. Medanic, "On stabilization and optimization by output feedback," in *Proc. 12th Asilomar Conf. Circ., Syst. Comp.*, Pacific Grove, CA, Nov. 1978, pp. 412-416.
- [5] J. V. Medanic, "Asymptotic properties of dynamic controllers designed by projective controls," in *Proc. 24th IEEE Conf. Decision Contr.*, Fort Lauderdale, FL, Dec. 1985.
- [6] D. S. Bernstein, "The optimal projection equations for static and dynamic output feedback: The singular case," *IEEE Trans. Automat. Contr.*, to be published.
- [7] S. Basuthakur and C. H. Knapp, "Optimal constant controllers for stochastic linear systems," *IEEE Trans. Automat. Contr.*, vol. AC-20, pp. 664-666, 1975.
- [8] W. M. Wonham, "Optimal stationary control of a linear system with state-dependent noise," *SIAM J. Contr.*, vol. 5, pp. 486-500, 1967.
- [9] D. Kleinman, "Optimal stationary control of linear systems with control-dependent noise," *IEEE Trans. Automat. Contr.*, vol. AC-14, pp. 673-677, 1969.
- [10] P. McLane, "Optimal stochastic control of linear systems with state and control-dependent disturbances," *IEEE Trans. Automat. Contr.*, vol. AC-16, pp. 793-798, 1971.
- [11] J. Bismut, "Linear-quadratic optimal stochastic control with random coefficients," *SIAM J. Contr.*, vol. 14, pp. 419-444, 1976.
- [12] D. S. Bernstein and D. C. Hyland, "The optimal projection/maximum entropy approach to designing low-order, robust controllers for flexible structures," in *Proc. 24th IEEE Conf. Decision Contr.*, Fort Lauderdale, FL, Dec. 1985, pp. 745-752.
- [13] D. S. Bernstein and S. W. Greeley, "Robust controller synthesis using the maximum entropy design equations," *IEEE Trans. Automat. Contr.*, vol. AC-33, pp. 362-364, 1986.
- [14] —, "Robust output-feedback stabilization: Deterministic and stochastic perspectives," in *Proc. Amer. Contr. Conf.*, Seattle, WA, June 1986, pp. 1818-1821.
- [15] L. Arnold, *Stochastic Differential Equations: Theory and Applications*. New York: Wiley, 1974.
- [16] D. C. Hyland and D. S. Bernstein, "The optimal projection equations for fixed-order dynamic compensation," *IEEE Trans. Automat. Contr.*, vol. AC-29, pp. 1034-1037, 1984.

Analysis of Time-Varying Scaled Systems Via General Orthogonal Polynomials

TSU TIAN LEE AND YIH FONG CHANG

Abstract—General orthogonal polynomials are introduced to analyze and approximate the solution of a class of scaled systems. Using the operational matrix of integration, together with the operational matrix of linear transformation, the dynamical equation of a scaled system is reduced to a set of simultaneous linear algebraic equations. The coefficient vectors of the general orthogonal polynomials can be determined recursively by the derived algorithm. An illustrative example is given to demonstrate the validity and applicability of the orthogonal polynomial approximations.

I. INTRODUCTION

An investigation of the dynamics of an overhead current collection mechanism for an electric locomotive by Ockendon and Taylor [12] revealed that under certain conditions, the dynamics of the systems is characterized by a differential equation containing terms with a scaled argument of the form

$$\dot{X}(t) = AX(\lambda t) + BX(t)$$

$$X(0) = X_0$$

where $X(\lambda t)$ and $X(t)$ are n -vectors and A and B are $n \times n$ matrices and the constant $0 < \lambda < 1$. This type of differential equation also plays an important role in several chemical processes [3], [13]. This equation was first studied by Fox *et al.* [11] with the introduction of a finite difference method for $0 < \lambda < 1$. Recently, the solution of such a scaled system has been obtained by several different orthogonal functions, such as block-pulse functions [14], [2], [3], Walsh functions [1], delayed unit step functions [4], Laguerre polynomials [5], Chebyshev polynomials [6], [7], and Legendre polynomials [15]. The common approach of these methods is the use of the operational matrix of integration together with the operational matrix of scaling to reduce the differential equation to a set of linear algebraic equations, which is more suitable for computer programming.

In this note we will employ the operational matrix of integration and product operational matrix of the general orthogonal polynomials, together with the operational matrix of linear transformation, which will be derived later, to obtain the solution of the scaled system. The operational matrix of linear transformation is derived based on the following properties, namely, the pure recurrence relation

$$\phi_{i+1}(z) = (a_i z + b_i)\phi_i(z) - c_i\phi_{i-1}(z) \quad (1)$$

with

$$\phi_0(z) = 1; \quad \phi_1(z) = a_0 z + b_0$$

and the differential recurrence relation

$$\dot{\phi}_i(z) = A_i\phi_{i+1}(z) + B_i\phi_i(z) + C_i\phi_{i-1}(z) \quad (2)$$

where recurrence coefficients a_i , b_i , c_i and differential recurrence coefficients A_i , B_i , and C_i , are specified by the particular orthogonal polynomials under consideration and some are listed in [9]. The aim of this paper is twofold: 1) to derive an operational matrix of linear transformation for general orthogonal polynomials so that the scaled

Manuscript received December 6, 1985; revised April 14, 1986 and September 9, 1986.

T. T. Lee is with the Department of Electrical Engineering, University of Kentucky, Lexington, KY 40506-0046.

Y. F. Chang is with the Institute of Electronics, National Chiao Tung University, Taiwan, Republic of China.

IEEE Log Number 8714806.

Optimal output feedback for non-zero set point regulation: the discrete-time case

WASSIM M. HADDAD† and DENNIS S. BERNSTEIN‡

Optimal discrete-time static output feedback is considered for a non-zero set point problem with non-zero mean disturbances. The optimal control law consists of a closed-loop component for feeding back the measurements and a constant open-loop component which accounts for the non-zero set point and non-zero disturbance mean. An additional feature is the presence of state-, control- and measurement-dependent white noise. It is shown that in the absence of multiplicative disturbances, the closed-loop controller can be designed independently of the open-loop control.

Notation and definitions

| | |
|-----------------------------|--|
| $R, R^{r \times s}, R^r, E$ | real numbers, $r \times s$ real matrices, $R^{r \times 1}$, expectation |
| $I_n, ()^T$ | $n \times n$ identity, transpose |
| \otimes | Kronecker product |
| $\text{tr } Z$ | trace of square matrix Z |
| asymptotically | |
| stable matrix | matrix with eigenvalues in the open unit disk |
| n, m, l, p | positive integers |
| x | n -dimensional vector |
| u, y | m -, l -dimensional vectors |
| $A, A_i; B, B_i; C, C_i$ | $n \times n$ matrices, $n \times m$ matrices, $l \times n$ matrices, $i = 1, \dots, p$ |
| L, K | $r \times n$ matrix, $m \times l$ matrix |
| δ, γ, α | r -, n -, m -dimensional vectors |
| k | discrete-time index $1, 2, \dots$ |
| $v_i(k)$ | unit variance white noise, $i = 1, \dots, p$ |
| $w_1(k), w_2(k)$ | n -dimensional, l -dimensional white noise processes |
| V_1, V_2 | $n \times n$ covariance of w_1 , $l \times l$ covariance of w_2 ; $V_1 \geq 0, V_2 \geq 0$ |
| V_{12} | $n \times l$ cross-covariance of w_1, w_2 |
| R_1, R_2 | $r \times r$ and $m \times m$ state and control weightings; $R_1 \geq 0, R_2 \geq 0$ |
| R_{12} | $r \times m$ cross weighting; $R_1 - R_{12}R_2^{-1}R_{12}^T \geq 0$ |
| \tilde{A}, \tilde{A}_i | $A + BKC, A_i + B_iKC + BKC_i, i = 1, \dots, p$ |
| \tilde{A} | $I_n - \tilde{A}$ |
| \tilde{B} | $B\alpha + \gamma$ |
| \tilde{B}_i | $B_i\alpha$ |
| \tilde{w} | $w_1 + BKw_2 + \sum_{i=1}^p B_iKw_2$ |

Received 9 March 1987.

† Department of Mechanical Engineering, Florida Institute of Technology, Melbourne, FL 32901, U.S.A.

‡ Harris Corporation, Government Aerospace Systems Division, MS 22/4848, Melbourne, FL 32902, U.S.A.

$$\begin{aligned}\tilde{V} &= V_1 + V_{12}K^T B^T + BK V_{12}^T + BK V_2 K^T B^T + \sum_{i=1}^p B_i K V_2 K^T B_i^T \\ \tilde{R} &= L^T R_1 L + L^T R_{12} K C + C^T K^T R_{12}^T L + C^T K^T R_2 K C \\ &\quad + \sum_{i=1}^p C_i^T K^T R_2 K C_i\end{aligned}$$

For arbitrary $m \in \mathbb{R}^n$ and $Q, P \in \mathbb{R}^{n \times n}$ define:

$$\begin{aligned}R_{2s} &\triangleq R_2 + B^T P B + \sum_{i=1}^p B_i^T P B_i, & V_{2s} &\triangleq V_2 + C Q C^T + \sum_{i=1}^p C_i (Q + m m^T) C_i^T \\ R_{2a} &\triangleq R_2 + B^T P B, & V_{2a} &\triangleq V_2 + C Q C^T \\ P_s &\triangleq B^T P A + R_{12}^T + \sum_{i=1}^p B_i^T P A_i, & Q_s &\triangleq A Q C^T + V_{12} + \sum_{i=1}^p A_i (Q + m m^T) C_i^T \\ P_a &\triangleq B^T P A + R_{12}^T, & Q_a &\triangleq A Q C^T + V_{12} \\ P_{s1} &\triangleq R_{12}^T + \sum_{i=1}^p B_i^T P A_i, & Q_{s1} &\triangleq V_{12} + \sum_{i=1}^p A_i (Q + m m^T) C_i^T\end{aligned}$$

1. Introduction

The quadratic performance criterion

$$J = \sum_{k=0}^N x^T(k) R_1 x(k) + u^T(k) R_2 u(k) \quad (1)$$

expresses the desire to minimize deviations of the state $x(k)$ of the system

$$x(k+1) = A x(k) + B u(k) + w(k) \quad (2)$$

from the regulation point $x=0$. As is well known (Kwakernaak and Sivan, 1972, pp. 504–509), the non-zero set point criterion

$$J_s = \sum_{k=0}^N [x(k) - \bar{x}]^T R_1 [x(k) - \bar{x}] + u^T(k) R_2 u(k) \quad (3)$$

presents no additional difficulty so long as $x(k)$ and $u(k)$ are replaced by $x(k) - \bar{x}$ and $u(k) - \bar{u}$, where \bar{u} satisfies

$$\bar{x} = A \bar{x} + B \bar{u} \quad (4)$$

Closer inspection, however, reveals that this approach is suboptimal. Specifically, the offset \bar{u} in the control may correspond to an unacceptably high level of control effort when $\bar{u}^T R_2 \bar{u}$ is large. Hence (3) overlooks design tradeoffs concerning the control effort required for maintaining the non-zero regulation point \bar{x} . Moreover, such an approach is impossible when \bar{u} satisfying (4) does not exist.

A significant advance in extending the full-state-feedback LQR formulation to steady-state periodic tracking problems (and hence to the special case of non-zero set point regulation) was given by Artstein and Leizarowitz (1985). Bernstein and Haddad (1987 b) generalize the results of Artstein and Leizarowitz (1985) for the non-zero set point regulation problem to include noisy and non-noisy measurements, weighted and unweighted controls, correlated plant/measurement noise, cross weighting, non-zero mean disturbances, and state-, control- and measurement-dependent multiplicative white noise. They consider the steady-state performance criterion

$$J = \lim_{t \rightarrow \infty} \mathbb{E}[(Lx(t) - \delta)^T R_1 (Lx(t) - \delta) + 2(Lx(t) - \delta)^T R_{12} u(t) + u^T(t) R_2 u(t)] \quad (5)$$

where δ is the non-zero regulation point. For full-state feedback with $R_{12} = 0$ and $L = \text{identity}$, Artstein and Leizarowitz (1985) show that for a constant offset control law

$$u(t) = Kx(t) + \alpha \quad (6)$$

K and α are given by

$$K = -R_2^{-1} B^T P \quad (7)$$

$$\alpha = -R_2^{-1} B^T (A - \Sigma P)^{-T} R_1 \delta \quad (8)$$

where P satisfies the Riccati equation

$$0 = A^T P + PA + R_1 - P \Sigma P$$

with

$$\Sigma \triangleq B R_2^{-1} B^T$$

Two features of the control law (6)–(8) are noteworthy. First, (6) consists of both a closed-loop feedback component $Kx(t)$ and an open-loop component α depending upon the regulation point δ . And, second (and more important), is the observation that the closed-loop control component is independent of the open-loop component. From a practical point of view this feature is quite useful since it implies that the feedback gain K can be determined without regard to the set point. Hence a change in the desired set point δ during on-line operation does not necessitate re-solving the Riccati equation in real time; only α requires updating. For a new value of δ , α can readily be recomputed on-line via the matrix multiplication operation (8). In the presence of multiplicative disturbances, however, the independence of the closed-loop component from the open-loop component is lost.

The purpose of the present paper is to provide a self-contained derivation of the optimality conditions for the non-zero set point problem in the discrete-time case. To obtain a realistic problem setting, we consider the case in which the full state is not available, but rather only noise-corrupted measurements of linear combinations of states. For greater design flexibility, we also allow the possibility for correlated plant and measurement noise. In addition, we consider the dual design feature, namely, cross weighting in the performance criterion. The presence of a non-zero constant plant disturbance in conjunction with zero-mean white plant disturbances, i.e. a non-zero mean disturbance, is also considered. Our results show that the presence of a non-zero constant disturbance component leads to an additional offset in the open-loop component of the control. Finally, in addition to the above generalizations we allow for the presence of multiplicative disturbances in the plant. The control law thus generalizes previous results involving state-, control- and measurement-dependent noise (Bernstein and Haddad 1987). As shown in Bernstein and Greeley (1986) and Haddad (1987), the multiplicative white noise model can be used for robustness with respect to plant parameter variations.

2. Non-zero set point regulation

2.1. Non-zero set point problem

Given the n th-order controlled system

$$x(k+1) = \left(A + \sum_{i=1}^p v_i(k) A_i \right) x(k) + \left(B + \sum_{i=1}^p v_i(k) B_i \right) u(k) + w_1(k) + \gamma \quad (9)$$

with measurements

$$y(k) = \left(C + \sum_{i=1}^p v_i(k) C_i \right) x(k) + w_2(k) \quad (10)$$

where $k = 1, 2, \dots$, determine K and α such that the static output feedback controller

$$u(k) = Ky(k) + \alpha \quad (11)$$

minimizes the steady-state performance criterion

$$J(K, \alpha) \triangleq \lim_{k \rightarrow \infty} E[(Lx(k) - \delta)^T R_1 (Lx(k) - \delta) + 2(Lx(k) - \delta)^T R_{12} u(k) + u^T(k) R_2 u(k)] \quad (12)$$

Using the notation of § 1 the closed-loop system (9)–(11) can be written as

$$x(k+1) = \left(\tilde{A} + \sum_{i=1}^p v_i(k) \tilde{A}_i \right) x(k) + \tilde{B} + \sum_{i=1}^p v_i(k) \tilde{B}_i + \tilde{w}(k) \quad (13)$$

To analyse (13) define the second-moment and covariance matrices

$$\tilde{Q}(k) \triangleq E[x(k)x^T(k)], \quad Q(k) \triangleq \tilde{Q}(k) - m(k)m^T(k)$$

where $m(k) \triangleq E[x(k)]$. It follows from (13) that $\tilde{Q}(k)$, $Q(k)$ and $m(k)$ satisfy

$$\begin{aligned} \tilde{Q}(k+1) &= \tilde{A}\tilde{Q}(k)\tilde{A}^T + \tilde{A}m(k)\tilde{B}^T + \tilde{B}m^T(k)\tilde{A}^T + \tilde{B}\tilde{B}^T \\ &\quad + \sum_{i=1}^p [\tilde{A}_i\tilde{Q}(k)\tilde{A}_i^T + \tilde{B}_i m^T(k)\tilde{A}_i^T + \tilde{A}_i m(k)\tilde{B}_i^T + \tilde{B}_i\tilde{B}_i^T] + \tilde{V} \end{aligned} \quad (14)$$

$$\begin{aligned} Q(k+1) &= \tilde{A}Q(k)\tilde{A}^T + \sum_{i=1}^p [\tilde{A}_iQ(k)\tilde{A}_i^T + \tilde{A}_i m(k)m^T(k)\tilde{A}_i^T \\ &\quad + \tilde{B}_i m^T(k)\tilde{A}_i^T + \tilde{A}_i m(k)\tilde{B}_i^T] + \tilde{V} \end{aligned} \quad (15)$$

$$m(k+1) = \tilde{A}m(k) + \tilde{B} \quad (16)$$

To consider the steady state, we restrict our consideration to the set of closed-loop second-moment stabilizing gains

$$\mathbf{S}_s \triangleq \left\{ K : \tilde{A} \otimes \tilde{A} + \sum_{i=1}^p \tilde{A}_i \otimes \tilde{A}_i \text{ is asymptotically stable} \right\}$$

It follows from fundamental properties of Lyapunov equations that if $K \in \mathbf{S}_s$, then \tilde{A} is also asymptotically stable. Hence, for $K \in \mathbf{S}_s$, $\tilde{Q} \triangleq \lim_{k \rightarrow \infty} \tilde{Q}(k)$, $Q \triangleq \lim_{k \rightarrow \infty} Q(k)$ and $m \triangleq \lim_{k \rightarrow \infty} m(k)$ exist and satisfy

$$\begin{aligned} \tilde{Q} &= \tilde{A}\tilde{Q}\tilde{A}^T + \tilde{A}m\tilde{B}^T + \tilde{B}m^T\tilde{A}^T + \tilde{B}\tilde{B}^T \\ &\quad + \sum_{i=1}^p [\tilde{A}_i\tilde{Q}\tilde{A}_i^T + \tilde{B}_i m^T\tilde{A}_i^T + \tilde{A}_i m\tilde{B}_i^T + \tilde{B}_i\tilde{B}_i^T] + \tilde{V} \end{aligned} \quad (17)$$

$$Q = \tilde{A}Q\tilde{A}^T + \sum_{i=1}^p [\tilde{A}_iQ\tilde{A}_i^T + \tilde{A}_i m m^T \tilde{A}_i^T + \tilde{B}_i m^T \tilde{A}_i^T + \tilde{A}_i m \tilde{B}_i^T + \tilde{B}_i \tilde{B}_i^T] + \tilde{V} \quad (18)$$

$$m = \mathbf{A}^{-1} \tilde{B} \quad (19)$$

Note that since \tilde{A} is asymptotically stable, the inverse in (19) exists. For $K \in \mathbf{S}_s$, it now

follows that $J(K, \alpha)$ is given by

$$J(K, \alpha) = \text{tr} [(Q + mm^T)\tilde{R}] + \text{tr} [K^T R_2 K V_2] + \delta^T R_1 \delta - 2m^T L^T R_1 L \delta \\ + 2m^T L^T R_{12} \alpha - 2\delta^T R_{12} K C m - 2\delta^T R_{12} \alpha + 2m^T C^T K^T R_2 \alpha + \alpha^T R_2 \alpha \quad (20)$$

Associated with Q is its dual $P \geq 0$ which is the unique solution of

$$P = \tilde{A}^T P \tilde{A} + \sum_{i=1}^p \tilde{A}_i^T P \tilde{A}_i + \tilde{R} \quad (21)$$

To obtain closed-form expressions for the feedback gain K , we further restrict consideration to the set

$$\mathbf{S}_s^+ \triangleq \{K \in \mathbf{S}_s : R_{2s} > 0, V_{2s} > 0 \text{ and } \Psi_s \text{ is invertible}\}$$

where

$$\Psi_s \triangleq B^T A^{-T} L^T R_1 L A^{-1} B + B^T A^{-T} L^T R_{12} (I_m + K C A^{-1} B) \\ + (I_m + K C A^{-1} B)^T R_{12}^T L A^{-1} B + (I_m + K C A^{-1} B)^T R_2 (I_m + K C A^{-1} B) \\ + \sum_{i=1}^p [B^T A^{-T} \tilde{A}_i^T P \tilde{A}_i A^{-1} B + B_i^T P \tilde{A}_i A^{-1} B + B^T A^{-T} A_i^T P B_i \\ + B_i^T P B_i + B^T A^{-T} C_i^T K^T R_2 K C_i A^{-1} B]$$

Furthermore, we assume that

$$[B_i \neq 0 \Rightarrow C_i = 0], \quad i = 1, \dots, p \quad (22)$$

i.e. for each $i \in \{1, \dots, p\}$, B_i and C_i are not both non-zero. Essentially, (22) expresses the condition that the control-dependent and measurement-dependent disturbances are independent. There are no constraints, however, on correlation with the state-dependent noise. For the statement of the main theorem define

$$\Lambda_s \triangleq B^T A^{-T} L^T (R_1 L + R_{12} K C) A^{-1} + (I_m + K C A^{-1} B)^T (R_{12}^T L + R_2 K C) A^{-1} \\ + \sum_{i=1}^p [(A_i A^{-1} B + B_i)^T P \tilde{A}_i A^{-1} + B^T A^{-T} C_i^T K^T R_2 K C_i A^{-1}] \\ \Omega \triangleq B^T A^{-T} L^T R_1 L + R_{12}^T L + B^T A^{-T} C^T K^T R_{12}^T L$$

Theorem 2.1

Suppose K and α solve the non-zero set point problem with $K \in \mathbf{S}_s^+$. Then there exist $n \times n$ Q , $P \geq 0$ such that

$$K = -R_{2s}^{-1} (B^T P A Q C^T + P_{s1} Q C^T + B^T P Q_{s1}) V_{2s}^{-1} \quad (23)$$

$$\alpha = -\Psi_s^{-1} [\Lambda_s \gamma + \Omega \delta] \quad (24)$$

and such that Q and P satisfy

$$Q = A Q A^T + V_1 + \sum_{i=1}^p [(A_i + B_i K C) Q (A_i + B_i K C)^T + B_i K V_2 K^T B_i^T + \tilde{A}_i m m^T \tilde{A}_i^T \\ + \tilde{B} m^T \tilde{A}_i^T + \tilde{A}_i m \tilde{B}^T + \tilde{B}_i \tilde{B}_i^T] \\ + (Q_s + B K V_{2s}) V_{2s}^{-1} (Q_s + B K V_{2s})^T - Q_s V_{2s} Q_s^T \quad (25)$$

$$P = A^T P A + R_1 + \sum_{i=1}^p [(A_i + B K C_i)^T P (A_i + B K C_i) + C_i^T K^T R_2 K C_i \\ + (P_s + R_{2s} K C)^T R_{2s}^{-1} (P_s + R_{2s} K C) - P_s^T R_{2s} P_s] \quad (26)$$

Proof

The derivation of the necessary conditions is a straightforward application of the Lagrange multiplier technique. To optimize (20) over \mathbf{S}_s^+ subject to the constraints (18) and (19), form the lagrangian

$$\mathbf{L}(K, \alpha, Q, P, m) \triangleq \text{tr} \left[\lambda_0 J(K, \alpha) + \left(\tilde{A}Q\tilde{A}^T + \sum_{i=1}^p [\tilde{A}_i Q \tilde{A}_i^T + \tilde{A}_i m m^T \tilde{A}_i^T + \tilde{B}_i m^T \tilde{A}_i^T + \tilde{A}_i m \tilde{B}_i^T + \tilde{B}_i \tilde{B}_i^T] + \tilde{V} - Q \right) P + \lambda^T (\tilde{A}m + \tilde{B} - m) \right]$$

where the Lagrange multipliers $\lambda_0 \geq 0$, $\lambda \in \mathbf{R}^n$ and $P \in \mathbf{R}^{n \times n}$ are not all zero. Setting $\partial \mathbf{L} / \partial Q = 0$ and using the second-moment stability assumption it follows that $\lambda_0 = 1$ without loss of generality. Thus the stationarity conditions are given by

$$\frac{\partial \mathbf{L}}{\partial Q} = \tilde{A}^T P \tilde{A} + \sum_{i=1}^p \tilde{A}_i^T P \tilde{A}_i + \tilde{R} - P = 0 \quad (27)$$

$$\begin{aligned} \frac{\partial \mathbf{L}}{\partial P} &= \tilde{A}Q\tilde{A}^T + \sum_{i=1}^p [\tilde{A}_i Q \tilde{A}_i^T + \tilde{A}_i m m^T \tilde{A}_i^T + \tilde{B}_i m^T \tilde{A}_i^T + \tilde{A}_i m \tilde{B}_i^T + \tilde{B}_i \tilde{B}_i^T] \\ &\quad + \tilde{V} - Q = 0 \end{aligned} \quad (28)$$

$$\frac{\partial \mathbf{L}}{\partial \lambda^T} = \tilde{A}m + \tilde{B} - m = 0 \quad (29)$$

$$\begin{aligned} \frac{\partial \mathbf{L}}{\partial K} &= R_{2s} K V_{2s} + R_{12}^T Q C^T + B^T P A Q C^T \\ &\quad + \sum_{i=1}^p [B_i P A_i Q C + B^T P A_i (Q + m m^T) C_i^T] + B^T P V_{12} = 0 \end{aligned} \quad (30)$$

$$\begin{aligned} \frac{\partial \mathbf{L}}{\partial \alpha} &= \sum_{i=1}^p [B_i^T P A_i m + B_i^T P B_i K C + B_i^T P B_i x] + \frac{1}{2} B^T \lambda + R_{12}^T L m \\ &\quad - R_{12}^T \delta + R_2 K C m + R_2 \alpha = 0 \end{aligned} \quad (31)$$

$$\begin{aligned} \frac{\partial \mathbf{L}}{\partial m} &= \tilde{R}m + \sum_{i=1}^p [\tilde{A}_i^T P B_i + \tilde{A}_i^T P \tilde{A}_i m] - \frac{1}{2} \mathbf{A}^T \lambda - L^T R_1 \delta + L^T R_{12} x \\ &\quad - C^T K^T R_{12}^T \delta + C^T K^T R_2 x = 0 \end{aligned} \quad (32)$$

$$\begin{aligned} \lambda &= 2\mathbf{A}^{-T} \left(\tilde{R}m + \sum_{i=1}^p [\tilde{A}_i^T P \tilde{B}_i + \tilde{A}_i P \tilde{A}_i^T m] - L^T R_1 \delta + L^T R_{12} x \right. \\ &\quad \left. - C^T K^T R_{12}^T \delta + C^T K^T R_2 x \right) \end{aligned} \quad (33)$$

Using the definitions for Q_{s1} and P_{s1} along with (33), we obtain (23) and (24). Substituting the expressions for the optimal gains into (27) and (28) yields (25) and (26). \square

Remark 1

Because of the presence of δ in (25) via m in both Q_{s1} and V_{2s} and in (25) via \tilde{B} (in m) and \tilde{B}_i , the closed-loop component of the control law (23) cannot be

determined independently of the open-loop component. As shown in the following section, independence is recovered when the multiplicative noise terms are absent.

Remark 2

To specialize Theorem 2.1 to the standard regulation problem, set $\delta = 0$ and $\gamma = 0$ yielding Theorem 2.1 of Bernstein and Haddad (1987 a).

3. Specializations of Theorem 2.1

A series of specializations of Theorem 2.1 is now given. We begin by deleting all multiplicative white noise terms, i.e.

$$A_i, B_i, C_i = 0, \quad i = 1, \dots, p \quad (34)$$

In this case the stabilizing set \mathbf{S}_s can be characterized by

$$\mathbf{S} = \{K: \tilde{A} \text{ is asymptotically stable}\}$$

and, furthermore, \mathbf{S}_s^+ becomes

$$\mathbf{S}^+ \triangleq \{K \in \mathbf{S}: R_{2a} > 0, V_{2a} > 0 \text{ and } \Psi_a \text{ is invertible}\}$$

where

$$\begin{aligned} \Psi_a \triangleq & B^T A^{-T} L^T R_1 L A^{-1} B + B^T A^{-T} L^T R_{12} (I_m + K C A^{-1} B) \\ & + (I_m + K C A^{-1} B)^T R_{12}^T L A^{-1} B + (I_m + K C A^{-1} B)^T R_2 (I_m + K C A^{-1} B) \end{aligned}$$

For the statement of Corollary 3.1 define

$$\Lambda_a \triangleq B^T A^{-T} L^T (R_1 L + R_{12} K C) A^{-1} + (I_m + K C A^{-1} B)^T (R_{12}^T L + R_2 K C) A^{-1}$$

Corollary 3.1

Assume (34) is satisfied and suppose K and α solve the non-zero set point problem with $K \in \mathbf{S}^+$. Then there exist $n \times n$ $Q, P \geq 0$ such that

$$K = -R_{2a}^{-1} (B^T P A Q C^T + R_{12}^T Q C^T + B^T P V_{12}) V_{2a}^{-1} \quad (35)$$

$$\alpha = -\Psi_a^{-1} [\Lambda_a \gamma + \Omega \delta] \quad (36)$$

and such that Q and P satisfy

$$Q = A Q A^T + V_1 + (Q_a + B K V_{2a}) V_{2a}^{-1} (Q_a + B K V_{2a})^T - Q_a V_{2a} Q_a^T \quad (37)$$

$$P = A^T P A + R_1 + (P_a + R_{2a} K C)^T R_{2a}^{-1} (P_a + R_{2a} K C) - P_a^T R_{2a} P_a \quad (38)$$

Finally, setting

$$\gamma = 0, \quad R_{12} = 0, \quad V_{12} = 0, \quad r = n, \quad L = I_n \quad (39)$$

we obtain the discrete-time version of Artstein and Leizarowitz (1985) for the case of output feedback. Define

$$\mathbf{S}_1^+ \triangleq \{K \in \mathbf{S}: R_{2a} > 0, V_{2a} > 0 \text{ and } \Psi_1 > 0\}$$

where

$$\Psi_1 \triangleq B^T A^{-T} R_1 A^{-1} B + (I_m + K C A^{-1} B)^T R_2 (I_m + K C A^{-1} B)$$

Corollary 3.2

Assume (34) and (39) are satisfied and suppose K and α solve the non-zero set

point problem with $K \in \mathbf{S}_1^+$. Then there exist $n \times n$ $Q, P \geq 0$ such that

$$K = -R_{2a}^{-1} B^T P A Q C^T V_{2a}^{-1} \quad (40)$$

$$\alpha = -\Psi_1^{-1} B^T A^{-T} R_1 \delta \quad (41)$$

and such that Q and P satisfy

$$\begin{aligned} Q = & A Q A^T + V_1 + (A Q C^T + B K V_{2a}) V_{2a}^{-1} (A Q C^T + B K V_{2a})^T \\ & - A Q C^T V_{2a} C Q A^T \end{aligned} \quad (42)$$

$$\begin{aligned} P = & A^T P A + R_1 + (B^T P A + R_{2a} K C)^T R_{2a}^{-1} (B^T P A + R_{2a} K C) \\ & - A^T P B R_{2a} B^T P A \end{aligned} \quad (43)$$

4. Directions for further research

The extension to fixed-order dynamic compensation for non-zero set point regulation appears possible using the approach of Hyland and Bernstein (1984) and Haddad (1987). A generalization of Theorem 2.1 to design periodic tracking controllers (either static or dynamic) via the parameter optimization approach is being developed.

ACKNOWLEDGMENT

Dr Bernstein was supported in part by the Air Office of Scientific Research under contract F49620-86-C-0002.

REFERENCES

- ARTSTEIN, Z., and LEIZAROWITZ, A., 1985, *I.E.E.E. Trans. autom. Control*, **30**, 1123.
- BERNSTEIN, D. S., and GREELEY, S. W., 1986, *I.E.E.E. Trans. autom. Control*, **31**, 362.
- BERNSTEIN, D. S., and HADDAD, W. M., 1987 a, *Int. J. Control*, **46**, 65; 1987 b, *I.E.E.E. Trans. autom. Control*, **32**, 641.
- HADDAD, W. M., 1987, Robust optimal projection control-system synthesis. Ph.D. dissertation. Department of Mechanical Engineering, Florida Institute of Technology, Melbourne, FL.
- HYLAND, D. C., and BERNSTEIN, D. S., 1984, *I.E.E.E. Trans. autom. Control*, **29**, 1034.
- KWAKERNAK, K., and SIVAN, R., 1972, *Linear Optimal Control Systems* (New York: Wiley-Interscience).

Optimal Nonzero Set Point Regulation Via Fixed-Order Dynamic Compensation

WASSIM M. HADDAD AND DENNIS S. BERNSTEIN

Abstract—Standard LQG control theory is generalized to a regulation problem involving specified nonzero set points for the state and control variables and nonzero-mean disturbances. For generality, the results are obtained for the problem of fixed-order (i.e., not necessarily full-order) dynamic compensation. When the state, control, and disturbance offsets are set to zero and the compensator order is set equal to the plant dimension, the standard LQG result is recovered. These results provide the dynamic counterpart for the nonzero set point regulation results obtained in [1] via static controllers.

I. INTRODUCTION

As discussed in [1], the standard quadratic performance criterion expresses the desire to maintain the state and control variables in the neighborhood of the origin. If regulation is desired about nonzero state and control offsets, then, in special cases, the set points can be translated to the origin and standard theory can be applied (see, e.g., [2, pp. 270–276]). In general, however, (see [1]) such a translation may either be suboptimal or impossible. The latter situation may occur, for example, if the number of state components with specified nonzero set points is greater than the number of controls, while the former is the case when the control offset is particularly costly.

Motivated by the work of Leizarowitz and Artstein [3], [4] on the more general problems of periodic and nonperiodic tracking, the nonzero set point problem was addressed in [1] for the case of static output-feedback controllers. The goal of the present note is to derive analogous results for the case of dynamic compensation considered by Leizarowitz in [5]. As in [1], the solution we obtain has the satisfying feature that the closed-loop dynamic-feedback-compensation gains are independent of the open-loop control components which arise from the state and control set points. Thus, if the state set point is changed during operation, then only the open-loop control components require updating. Consequently, there is no need to recalculate the closed-loop gains by solving Riccati equations in real time. The overall theory thus permits the treatment of step commands within standard LQG theory.

For generality the development herein incorporates several special features which provide additional flexibility in applications. These include: 1) constant disturbance vectors in addition to zero-mean additive plant and measurement noise (i.e., nonzero-mean disturbances); 2) correlated plant and measurement noise; 3) state/control performance cross-weighting; 4) arbitrary set points for selected linear combinations of the state and control variables (see L_1 and L_2 in the problem statement in Section III); and 5) fixed-order (i.e., full- or reduced-order) compensation. Because of the last feature, the results obtained in the present note also generalize the results of [6]. For clarity, we specialize the main result to the usual full-order LQG case.

II. NOTATION AND DEFINITIONS

$R, R^{-1}, R', \mathbb{E}$ Real numbers, $r \times s$ real matrices, R^{-1} , expectation.

Manuscript received September 15, 1987; revised November 23, 1987. This work was supported in part by the Air Force Office of Scientific Research under Contract F49620-86-C-0002.

W. M. Haddad is with the Department of Mechanical and Aerospace Engineering, Florida Institute of Technology, Melbourne, FL 32901.

D. S. Bernstein is with Harris Corporation, Government Aerospace Systems Division, Melbourne, FL 32902.

IEEE Log Number 8821570.

$I_n, ()^T, ()'$

asymptotically stable matrix

n, m, l, q, r, n_c

\tilde{n}

x, u, y, x_c, \tilde{x}

A, B, C, D

A_c, B_c, C_c

L_1, L_2

δ_1, δ_2

γ_1, γ_2

α, α_c

$\delta, \tilde{\gamma}, \tilde{\alpha}$

$w_1(t), w_2(t)$

V_1, V_2

V_{12}

$\tilde{w}(t), \tilde{V}$

R_1, R_2

R_{12}

\tilde{R}

\tilde{A}, \tilde{B}

m, m_c

\tilde{m}

\tilde{R}_1

$\tilde{R}_{12}, \tilde{R}_2, \tilde{R}_2'$

\tilde{N}, \tilde{S}

$n \times n$ identity, transpose, group generalized inverse.

Matrix with eigenvalues in open left-half plane.

Positive integers.

$n + n_c$.

n, m, l, n_c, \tilde{n} -dimensional vectors.

$n \times n, n \times m, l \times n, l \times m$ matrices.

$n_c \times n_c, n_c \times l, m \times n_c$ matrices.

$q \times n, r \times m$ matrices.

q, r -dimensional set point vectors.

n, l -dimensional constant disturbance vectors.

m, n_c -dimensional control vectors.

$$\begin{bmatrix} \delta_1 \\ \delta_2 \end{bmatrix}, \begin{bmatrix} \gamma_1 \\ B_c \gamma_2 \end{bmatrix}, \begin{bmatrix} \alpha \\ \alpha_c \end{bmatrix}.$$

n, l -dimensional zero-mean white noise processes.

Intensities of w_1, w_2 ; $V_1 \geq 0, V_2 > 0$.

$n \times l$ cross intensity of w_1, w_2 .

$$\begin{bmatrix} w_1(t) \\ B_c w_2(t) \end{bmatrix}, \begin{bmatrix} V_1 & V_{12} B_c^T \\ B_c V_{12}^T & B_c V_2 B_c^T \end{bmatrix}.$$

$q \times q$ and $r \times r$ state and control weightings;

$R_1 \geq 0, R_2 \geq 0, L_2^T R_2 L_2 > 0$.

$q \times r$ cross weighting; $L_1^T R_1 L_1 - L_1^T R_{12} L_2 (L_2^T R_2 L_2)^{-1} L_2^T R_{12}^T L_1 \geq 0$.

$$\begin{bmatrix} L_1^T R_1 L_1 & L_1^T R_{12} L_2 C_c \\ C_c^T L_1^T R_{12}^T L_1 & C_c^T L_2^T R_2 L_2 C_c \end{bmatrix}.$$

$$\begin{bmatrix} A & B C_c \\ B_c C & A_c + B_c D C_c \end{bmatrix}, \begin{bmatrix} B & 0 \\ B_c D & I_{n_c} \end{bmatrix}.$$

n, n_c -dimensional vectors.

$$\begin{bmatrix} m \\ m_c \end{bmatrix}.$$

$$\begin{bmatrix} L_1^T R_1 L_1 - L_1^T R_{12} L_2 (L_2^T R_2 L_2)^{-1} L_2^T R_{12}^T L_1 & 0 \\ 0 & 0 \end{bmatrix}.$$

$$\begin{bmatrix} L_2^T R_{12}^T L_1 & L_2^T R_2 L_2 C_c \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} L_2^T R_2 L_2 & 0 \\ 0 & 0 \end{bmatrix}.$$

$$\begin{bmatrix} (L_2^T R_2 L_2)^{-1} & 0 \\ 0 & 0 \end{bmatrix}.$$

$$\begin{bmatrix} L_2^T R_{12}^T L_1 & L_2^T R_2 L_2 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} L_1^T R_1 & L_1^T R_{12} \\ C_c^T L_1^T R_{12}^T L_1 & C_c^T L_2^T R_2 L_2 \end{bmatrix}.$$

For arbitrary $n \times n$ Q, P define:

$$Q_c \triangleq Q C^T + V_{12}, \quad P_c \triangleq B^T P + L_2^T R_{12}^T L_1,$$

$$A_Q \triangleq A - Q_c V_2^{-1} C^T, \quad A_P \triangleq A - B (L_2^T R_2 L_2)^{-1} P_c.$$

III. DYNAMIC COMPENSATION FOR NONZERO SET POINT REGULATION

A. Nonzero Set Point Problem

Given the n th-order stabilizable and detectable plant

$$\dot{x}(t) = Ax(t) + Bu(t) + w_1(t) + \gamma_1, \quad t \in [0, \infty), \quad (3.1)$$

$$y(t) = Cx(t) + Du(t) + w_2(t) + \gamma_2 \quad (3.2)$$

design a fixed-order dynamic compensator

$$\dot{x}_c(t) = A_c x_c(t) + B_c y(t) + \alpha_c, \quad (3.3)$$

$$u(t) = C_c x_c(t) + \alpha \quad (3.4)$$

which minimizes the steady-state performance criterion

$$\begin{aligned} J(A_c, B_c, C_c, \alpha, \alpha_c) \triangleq & \lim_{t \rightarrow \infty} \mathbb{E}[(L_1 x(t) - \delta_1)^T R_1 (L_1 x(t) - \delta_1) \\ & + 2(L_1 x(t) - \delta_1)^T R_{12} (L_2 u(t) - \delta_2) \\ & + (L_2 u(t) - \delta_2)^T R_2 (L_2 u(t) - \delta_2)]. \end{aligned} \quad (3.5)$$

Remark 3.1: The cost functional (3.5) is identical to the LQG criterion (usually stated in terms of an averaged integral) with the exception of the shifted set points δ_1 and δ_2 and matrices L_1 and L_2 for selecting linear combinations of components of x and u .

The closed-loop system (3.1)–(3.4) can be written as

$$\dot{\bar{x}}(t) = \bar{A}\bar{x}(t) + \bar{B}\bar{u}(t) + \bar{w}(t) + \bar{\gamma}, \quad t \in [0, \infty) \quad (3.6)$$

where $\bar{x}(t) \triangleq [x^T(t), x_c^T(t)]^T$ and the closed-loop disturbance $\bar{w}(t)$ has nonnegative-definite intensity \bar{V} . To analyze (3.6) define the covariance matrix

$$\bar{Q}(t) \triangleq \mathbb{E}[(\bar{x}(t) - \bar{m}(t))(\bar{x}(t) - \bar{m}(t))^T] = \mathbb{E}[\bar{x}(t)\bar{x}^T(t)] - \bar{m}(t)\bar{m}^T(t)$$

where $\bar{m}(t) \triangleq \mathbb{E}[\bar{x}(t)]$. As shown in [1], $\bar{Q}(t)$ and $\bar{m}(t)$ satisfy

$$\dot{\bar{Q}}(t) = \bar{A}\bar{Q}(t) + \bar{Q}(t)\bar{A}^T + \bar{V}, \quad (3.7)$$

$$\dot{\bar{m}}(t) = \bar{A}\bar{m}(t) + \bar{B}\bar{\alpha} + \bar{\gamma}. \quad (3.8)$$

To guarantee that J is finite and independent of initial conditions, we restrict our attention to the set of admissible stabilizing compensators

$$\mathcal{S} \triangleq \{(A_c, B_c, C_c): \bar{A} \text{ is asymptotically stable}\}.$$

Hence, for $(A_c, B_c, C_c) \in \mathcal{S}$, $\bar{Q} \triangleq \lim_{t \rightarrow \infty} \bar{Q}(t)$ and $\bar{m} \triangleq \lim_{t \rightarrow \infty} \bar{m}(t)$ exist and satisfy

$$0 = \bar{A}\bar{Q} + \bar{Q}\bar{A}^T + \bar{V}, \quad (3.9)$$

$$0 = \bar{A}\bar{m} + \bar{B}\bar{\alpha} + \bar{\gamma}. \quad (3.10)$$

Since the value of J is independent of the internal realization of the transfer function corresponding to (3.3) and (3.4), without loss of

generality we further restrict our attention to the set

$$\mathcal{S}' \triangleq \{(A_c, B_c, C_c) \in \mathcal{S}: (A_c, B_c) \text{ is controllable} \\ \text{and } (A_c, C_c) \text{ is observable}\}.$$

Now $J(A_c, B_c, C_c, \alpha, \alpha_c)$ is given by

$$\begin{aligned} J(A_c, B_c, C_c, \alpha, \alpha_c) = & \text{tr} [\bar{Q} + \bar{m}\bar{m}^T] \bar{R} - 2m^T L_1^T R_1 \delta_1 + \delta_1^T R_1 \delta_1 \\ & + 2m^T L_1^T R_{12} L_2 \alpha - 2m^T L_1^T R_{12} \delta_2 \\ & - 2\delta_1^T R_{12} L_2 C_c m_c - 2\delta_1^T R_{12} L_2 \alpha \\ & + 2\delta_1^T R_{12} \delta_2 + 2m_c^T C_c^T L_2^T R_2 L_2 \alpha - 2m_c^T C_c^T L_2^T R_2 \delta_2 \\ & - 2\alpha^T L_2^T R_2 \delta_2 + \alpha^T L_2^T R_2 L_2 \alpha + \delta_2^T R_2 \delta_2. \end{aligned} \quad (3.11)$$

To obtain closed-form expressions for the feedback gains we further restrict consideration to the set

$$\mathcal{S}'' \triangleq \{(A_c, B_c, C_c) \in \mathcal{S}': \Omega > 0\},$$

where

$$\Omega \triangleq \bar{B}^T \bar{A}^{-T} \bar{R}_1 \bar{A}^{-1} \bar{B} + (\bar{R}_{12} \bar{A}^{-1} \bar{B} - \bar{R}_{12})^T \bar{R}_2^{-1} (\bar{R}_{12} \bar{A}^{-1} \bar{B} - \bar{R}_{12}).$$

The following factorization lemma is needed for the statement of the main result.

Lemma 3.1: Suppose $n \times n$ \bar{Q} , \bar{P} are nonnegative definite and rank $\bar{Q}\bar{P} = n_c$. Then there exist $n_c \times n$ G , Γ and $n_c \times n_c$ invertible M such that

$$\bar{Q}\bar{P} = G^T M \Gamma, \quad (3.12)$$

$$\Gamma G^T = I_{n_c}. \quad (3.13)$$

Furthermore, G , M , and Γ are unique except for a change of basis in \mathbb{R}^{n_c} .

Proof: See [6]. \square

As shown in [6], $\bar{Q}\bar{P}$ has a group generalized inverse $(\bar{Q}\bar{P})^+ = G^T M^{-1} \Gamma$, and the matrix

$$\tau \triangleq \bar{Q}\bar{P}(\bar{Q}\bar{P})^+ = G^T \Gamma \quad (3.14)$$

is an oblique projection. A triple (G, M, Γ) satisfying (3.12) and (3.13) with $G, \Gamma \in \mathbb{R}^{n_c \times n}$, $M \in \mathbb{R}^{n_c \times n_c}$, and $n_c = \text{rank } \bar{Q}\bar{P}$ will be called a *projective factorization* of $\bar{Q}\bar{P}$. Furthermore, define the complementary projection $\tau_\perp \triangleq I_n - \tau$. Optimizing (3.11) subject to (3.9) and (3.10) yields the following result illustrated in Fig. 1.

Theorem 3.1: Suppose $(A_c, B_c, C_c, \alpha, \alpha_c)$ solves the nonzero set point problem with $(A_c, B_c, C_c) \in \mathcal{S}''$. Then there exist $n \times n$ nonnegative-definite matrices \bar{Q} , \bar{P} , \bar{Q} , \bar{P} such that, for some projective factorization (G, M, Γ) of $\bar{Q}\bar{P}$, A_c , B_c , C_c , α , and α_c are given by

$$A_c = \Gamma [A - B(L_1^T R_1 L_1)^{-1} P_c - Q_c V_c^{-1} C + Q_c V_c^{-1} D(L_1^T R_1 L_1)^{-1} P_c] G^T, \quad (3.15)$$

$$B_c = \Gamma Q_c V_c^{-1}, \quad (3.16)$$

$$C_c = -(L_1^T R_1 L_1)^{-1} P_c G^T, \quad (3.17)$$

$$\begin{bmatrix} \alpha \\ \alpha_c \end{bmatrix} = \Omega^{-1} [(\bar{R}_{12} - \bar{B}^T \bar{A}^{-T} \bar{R}) \bar{A}^{-1} \bar{\gamma} + (\bar{N} - \bar{B}^T \bar{A}^{-T} \bar{S}) \bar{\delta}] \quad (3.18)$$

$$\frac{\partial \mathcal{L}}{\partial \tilde{m}} = \tilde{R}\tilde{m} + \frac{1}{2} \tilde{A}^T \tilde{\lambda} + \tilde{R}_{12}^T \tilde{\alpha} - \tilde{S}\tilde{\delta} = 0, \quad (4.3)$$

$$\frac{\partial \mathcal{L}}{\partial \tilde{\alpha}} = \begin{bmatrix} [L_1^T R_1 L_2 \ 0] \tilde{\alpha} + [L_1^T R_{12}^T L_1 \ L_1^T R_2 L_2 C_c] \tilde{m} - [L_1^T R_{12} \ L_1^T R_2] \tilde{\delta} + \frac{1}{2} B^T \lambda_1 + \frac{1}{2} D^T B_c^T \lambda_2 \\ \frac{1}{2} \lambda_2 \end{bmatrix} = 0, \quad (4.4)$$

$$\frac{\partial \mathcal{L}}{\partial A_c} = P_{12}^T Q_{12} + P_2 Q_2 + \frac{1}{2} \lambda_2 m_c^T = 0, \quad (4.5)$$

$$\frac{\partial \mathcal{L}}{\partial B_c} = P_{12}^T V_{12} + P_2 B_c V_2 + (P_{12}^T Q_1 + P_2 Q_{12}^T) C^T + (P_{12}^T Q_{12} + P_2 Q_2) C_c^T D^T + \frac{1}{2} \lambda_2 m^T C^T + \frac{1}{2} \lambda_2 \gamma_2^T + \frac{1}{2} \lambda_2 \alpha^T D^T = 0, \quad (4.6)$$

$$\begin{aligned} \frac{\partial \mathcal{L}}{\partial C_c} = & L_1^T R_{12}^T L_1 Q_{12} + L_1^T R_2 L_2 C_c Q_2 + L_1^T R_{12}^T L_1 m m_c^T + L_1^T R_2 L_2 C_c m_c m_c^T + B^T (P_1 Q_{12} + P_{12} Q_2) + D^T B_c^T (P_{12}^T Q_{12} + P_2 Q_2) - L_1^T R_{12}^T \delta_1 m_c^T \\ & + L_1^T R_2 L_2 \alpha m_c^T - L_1^T R_2 \delta_2 m_c^T + \frac{1}{2} B^T \lambda_1 m_c^T = 0. \end{aligned} \quad (4.7)$$

Expanding (4.1) and (4.2) yields

$$0 = A Q_1 + Q_1 A^T + V_1 + B C_c Q_{12}^T + Q_{12} C_c^T B^T, \quad (4.8)$$

$$0 = A Q_{12} + Q_{12} A_c^T + B C_c Q_2 + Q_1 C^T B_c^T + V_{12} B_c^T + Q_{12} C_c^T D^T B_c^T, \quad (4.9)$$

$$0 = A_c Q_2 + Q_2 A_c^T + B_c C Q_{12} + Q_{12}^T C^T B_c^T + B_c V_2 B_c^T + B_c D C_c Q_2 + Q_2 C_c^T D^T B_c^T, \quad (4.10)$$

$$0 = A^T P_1 + P_1 A + L_1^T R_1 L_1 + C^T B_c^T P_{12}^T + P_{12} B_c C, \quad (4.11)$$

$$0 = A^T P_{12} + P_{12} A_c + C^T B_c^T P_2 + P_1 B C_c + L_1^T R_{12} L_2 C_c + P_{12} B_c D C_c, \quad (4.12)$$

$$0 = A_c^T P_2 + P_2 A_c + C_c^T B^T P_{12} + P_{12}^T B C_c + C_c^T L_1^T R_2 L_2 C_c + C_c^T D^T B_c^T P_2 + P_2 B_c D C_c. \quad (4.13)$$

Next, note that (4.4) implies that $\lambda_2 = 0$, and thus (4.5) can be written as

$$-P_2^{-1} P_{12}^T Q_{12} Q_2^{-1} = I_{n_c}. \quad (4.14)$$

The existence of Q_2^{-1} and P_2^{-1} follows from the fact that (A_c, B_c, C_c) is minimal. See [6] for details. Now define the $n \times n$ matrices

$$Q \triangleq Q_1 - Q_{12} Q_2^{-1} Q_{12}^T, \quad P \triangleq P_1 - P_{12} P_2^{-1} P_{12}^T,$$

$$\hat{Q} \triangleq Q_{12} Q_2^{-1} Q_{12}^T, \quad \hat{P} \triangleq P_{12} P_2^{-1} P_{12}^T,$$

$$\tau \triangleq -Q_{12} Q_2^{-1} P_2^{-1} P_{12}^T$$

and the $n_c \times n$, $n_c \times n_c$, and $n_c \times n$ matrices

$$G \triangleq Q_2^{-1} Q_{12}^T, \quad M \triangleq Q_2 P_2, \quad \Gamma \triangleq -P_2^{-1} P_{12}^T.$$

Note that $\tau = G^T \Gamma$. Clearly, Q , P , \hat{Q} , and \hat{P} are symmetric and nonnegative definite.

Next note that with the above definitions, (4.14) is equivalent to (3.13) and that (3.12) holds. Hence, $\tau = G^T \Gamma$ is idempotent, i.e., $\tau^2 = \tau$. Sylvester's inequality yields (3.23). Note also that

$$\hat{Q} = \tau \hat{Q}, \quad \hat{P} = \hat{P} \tau.$$

The components of \hat{Q} and \hat{P} can be written in terms of Q , P , \hat{Q} , \hat{P} , G , and Γ as

$$Q_1 = Q + \hat{Q}, \quad P_1 = P + \hat{P},$$

$$Q_{12} = \hat{Q} \Gamma^T, \quad P_{12} = -\hat{P} G^T,$$

$$Q_2 = \Gamma \hat{Q} \Gamma^T, \quad P_2 = G \hat{P} G^T.$$

The expressions (3.16) and (3.17) follow from (4.6) and (4.7) by using the n_c and n components of (4.4), respectively, and the above identities. Next, computing either $\Gamma(4.9)-(4.10)$ or $G(4.12) + (4.13)$ yields (3.15).

Substituting this expression for A_c into (4.8)-(4.13) it follows that (4.10) = $\Gamma(4.9)$ and (4.13) = $G(4.12)$. Thus, (4.10) and (4.13) are superfluous and can be omitted. Next, using (4.8) + $G^T \Gamma(4.9)G - (4.9)G - [(4.9)G]^T$ and $G^T \Gamma(4.9)G - [(4.9)G]^T$ yields (3.19) and (3.21). Using (4.11) + $\Gamma^T G(4.12)\Gamma - (4.12)\Gamma - [(4.12)\Gamma]^T$ and $\Gamma^T G(4.12)\Gamma - (4.12)\Gamma - [(4.12)\Gamma]^T$ yields (3.20) and (3.22).

To obtain (3.18) note that (4.4) can be rewritten as

$$\tilde{R}_2 \tilde{\alpha} + \tilde{R}_{12} \tilde{m} - \tilde{N} \tilde{\delta} + \frac{1}{2} \tilde{B}^T \tilde{\lambda} = 0. \quad (4.15)$$

Next, note that (4.3) is equivalent to

$$\frac{1}{2} \tilde{\lambda} = -\tilde{A}^{-T} \tilde{R} \tilde{m} + \tilde{A}^{-T} \tilde{S} \tilde{\delta} - \tilde{A}^{-T} \tilde{R}_{12}^T \tilde{\alpha}. \quad (4.16)$$

Substituting (4.16) into (4.15) yields

$$(\tilde{R}_2 - \tilde{B}^T \tilde{A}^{-T} \tilde{R}_{12}^T) \tilde{\alpha} + (\tilde{R}_{12} - \tilde{B}^T \tilde{A}^{-T} \tilde{R}) \tilde{m} + (\tilde{B}^T \tilde{A}^{-T} \tilde{S} - \tilde{N}) \tilde{\delta} = 0. \quad (4.17)$$

Next, note that (3.10) is equivalent to

$$\tilde{m} = -\tilde{A}^{-1} \tilde{B} \tilde{\alpha} - \tilde{A}^{-1} \tilde{\gamma}. \quad (4.18)$$

Now, substituting (4.18) into (4.17) yields

$$(\tilde{R}_2 - \tilde{R}_{12} \tilde{A}^{-1} \tilde{B} - \tilde{B}^T \tilde{A}^{-T} \tilde{R}_{12}^T + \tilde{B}^T \tilde{A}^{-T} \tilde{R} \tilde{A}^{-1} \tilde{B}) \tilde{\alpha} = (\tilde{N} - \tilde{B}^T \tilde{A}^{-T} \tilde{S}) \tilde{\delta} + (\tilde{R}_{12} - \tilde{B}^T \tilde{A}^{-T} \tilde{R}) \tilde{A}^{-1} \tilde{\gamma}. \quad (4.20)$$

Finally, note that the coefficient of $\tilde{\alpha}$ in (4.20) is equivalent to Ω and thus (4.20) yields (3.18). \square

V. CONCLUDING REMARKS

The results of the present note can be combined with the results of [1] to obtain nonstrictly proper controllers leading to a generalization of [7]. Current research is focused on extending the results of the present note to larger classes of command and disturbance signals.

REFERENCES

- [1] D. S. Bernstein and W. M. Haddad, "Optimal output feedback for nonzero set point regulation," *IEEE Trans. Automat. Contr.*, vol. AC-32, pp. 641-645, 1987.
- [2] H. Kwakernaak and R. Sivan, *Linear Optimal Control Systems*. New York: Wiley, 1972.
- [3] Z. Artstein and A. Leizarowitz, "Tracking periodic signals with the overtaking criterion," *IEEE Trans. Automat. Contr.*, vol. AC-30, pp. 1123-1126, 1985.
- [4] A. Leizarowitz, "Tracking nonperiodic trajectories with the overtaking criterion," *Appl. Math. Optimiz.*, vol. 14, pp. 155-171, 1986.
- [5] —, "Infinite horizon stochastic regulation and tracking with the overtaking criterion," *Stochastics*, 1987.
- [6] D. C. Hyland and D. S. Bernstein, "The optimal projection equations for fixed-

order dynamic compensation," *IEEE Trans. Automat. Contr.*, vol. AC-29, pp. 1034-1037, 1984.

- [7] D. S. Bernstein, "The optimal projection equations for static and dynamic output feedback: The singular case," *IEEE Trans. Automat. Contr.*, vol. AC-32, pp. 1139-1143, 1987.
- [8] S. Richter, "A homotopy algorithm for solving the optimal projection equations for fixed-order dynamic compensation: Existence, convergence and global optimality," in *Proc. Amer. Contr. Conf.*, Minneapolis, MN, June 1987, pp. 1527-1531.

A Study of Controllability and Time-Optimal Control of a Robot Model with Drive Train Compliances and Actuator Dynamics

A. AILON AND G. LANGHOLZ

Abstract—The problems of robot controllability and time-optimal control where drive train compliances and actuator dynamics are incorporated in the mathematical model is the subject of this note. This study demonstrates the conditions that ensure the existence of a time-optimal control, and establishes controllability of the augmented model (robot and actuator) in open- and closed-loop form. This note describes a procedure for the derivation of easily computable functional inequalities which represent upper bounds on the norm of the augmented system's time response.

I. INTRODUCTION

To obtain the control strategy of mechanical manipulators, various control schemes are presented in the available literature. A few examples are resolved control [1], inverse problems technique [2], and resolved acceleration control [3]. In most cases, the control scheme involves the computation of the appropriate generalized forces by the equation

$$H(\theta)\ddot{\theta}^{(2)} + K(\theta, \dot{\theta}) + R(\theta) = q$$

where θ and q are the vectors of the generalized coordinates and forces, respectively. H is the moment of inertia matrix, K is a vector specifying centrifugal and Coriolis effects, and R is a vector specifying gravitational effects.

In much of the literature the actuators providing the drive torques are modeled as pure torque sources. However, this approach is in most cases a simplification of the realistic models of the system [4]-[8].

The objective of this note is to study controllability and to investigate the conditions which ensure the existence of a control function that transfers the augmented model of the mechanical system, the actuator's dynamics, and the drive train's compliances, from a given initial position to a desired target in a minimum time. The model and the approach are useful for the design of a linear controller and can be used as a point of departure for a more general model of a robot arm.

II. THE MATHEMATICAL MODEL

The Lagrange formulation of a multilink mechanical system is given by

$$d(\partial L / \partial \dot{\theta}_i) / dt - \partial L / \partial \theta_i = q_i, \quad i = 1, 2, \dots, n \quad (1)$$

Manuscript received April 18, 1985; revised July 11, 1986, April 10, 1987, and September 23, 1987.

A. Ailon is with the Department of Electrical and Computer Engineering, Ben-Gurion University of the Negev, Beer Sheva, Israel.

G. Langholz is with the Department of Electrical Engineering, Florida State University, Tallahassee, FL 32301.

IEEE Log Number 8821571.

where $L = T - V$. T and V are the kinetic and potential energies of the system, respectively.

Let p_i be the i th generalized momentum [9]. Using Legendre's dual transformation

$$p_i = \partial L / \partial \dot{\theta}_i, \quad i = 1, 2, \dots, n. \quad (2)$$

Since L is a quadratic function in $\dot{\theta}$, p is linear in $\dot{\theta}$ for any given θ , i.e.,

$$p_i = \sum_{j=1}^n a_{ij}(\theta) \dot{\theta}_j, \quad i = 1, 2, \dots, n \quad (3)$$

with $a_{ij}(\theta) = \partial^2 L / \partial \dot{\theta}_i \partial \dot{\theta}_j$.

The inertial matrix is $H = [\partial^2 L / \partial \dot{\theta}_i \partial \dot{\theta}_j]_{n \times n}$ with $\det(H) = h(\theta) > 0$, $\forall \theta$, where $\det(\cdot)$ is the determinant of (\cdot) . Now, from (1), (2), and (3) we have

$$\dot{p}_i = \partial L / \partial \theta_i + q_i, \quad i = 1, 2, \dots, n, \quad (4)$$

$$\dot{\theta}_i = \sum_j b_{ij}(\theta) p_j, \quad i = 1, 2, \dots, n. \quad (5)$$

Using (5) one obtains

$$\partial L / \partial \theta_i = \left[\sum_{j=1}^n \sum_{k=1}^n c_{ijk}(\theta) p_j p_k \right] / (\det(H))^2. \quad (6)$$

Equations (4)-(6) constitute the state equations of the n -link mechanical system which can be written as

$$\dot{z}(t) = F(z(t)) + Bq(t), \quad z(t_0) = z_0 \quad (7)$$

where the vectors $z = [p^T \theta^T]^T$, $q = [q_1 q_2 \dots q_n]^T$, and $F = [F_1 F_2 \dots F_{2n}]^T$ are in Euclidean vector space with the usual norm $\|z\|^2 = \sum_{i=1}^{2n} (z_i)^2$. We also have

$$F_s = \left[\sum_{j=1}^n \sum_{k=1}^n c_{sjk} p_j p_k \right] / (\det(H))^2, \quad s = 1, 2, \dots, n$$

$$= \left[\sum_{i=1}^n d_{si} p_i \right] / \det(H), \quad s = n+1, n+2, \dots, 2n$$

and $B = \{I\}$, where I is the $n \times n$ identity matrix.

As an example, the exact equations for the two-link mechanical system which is confined to move in the vertical plane are given by

$$\begin{aligned} \dot{p}_1 &= [p_1 p_2 l_1 l_2 m_2 \sin(\theta_2 - \theta_1) \det(H) - \{0.5 p_1^2 I_2 + 0.5 p_2^2 (I_1 + m_2 l_1^2) \\ &\quad - p_1 p_2 E\} 2 E l_1 l_2 m_2 \sin(\theta_2 - \theta_1) / (\det(H))^2 \\ &\quad - (m_1 g l_1 + m_2 g l_1) \sin \theta_1 + q_1 = F_1(p_1, p_2, \theta_1, \theta_2) + q_1 \\ \dot{p}_2 &= -[p_1 p_2 l_1 l_2 m_2 \sin(\theta_2 - \theta_1) \det(H) - \{0.5 p_1^2 I_2 + 0.5 p_2^2 (I_1 + m_2 l_1^2) \\ &\quad - p_1 p_2 E\} 2 E l_1 l_2 m_2 \sin(\theta_2 - \theta_1) / (\det(H))^2 - m_2 g l_2 \sin \theta_2 + q_2 \\ &= F_2(p_1, p_2, \theta_1, \theta_2) + q_2 \\ \dot{\theta}_1 &= (p_1 I_2 - p_2 E) / \det(H) = F_3(p_1, p_2, \theta_1, \theta_2) \\ \dot{\theta}_2 &= (p_2 (I_1 + m_2 l_1^2) - p_1 E) / \det(H) = F_4(p_1, p_2, \theta_1, \theta_2) \end{aligned} \quad (8)$$

where $E = l_1 l_2 m_2 \cos(\theta_2 - \theta_1)$, m_i and l_i are the mass and the length of the i th link, respectively, I_i is the moment of inertia of the i th link with respect to the i th joint, and l_{ci} is the distance from the i th joint to the center of gravity of the i th link.

The term $\det(H)$ is a trigonometric function of, and periodical in, θ_i . This function attains its minimum in the interval $0 \leq \theta_i \leq 2\pi$, $i = 1, 2, \dots, n$, and therefore

$$\det(H) \geq k > 0, \quad \forall z \in R^{2n}. \quad (9)$$

We turn now to the dynamics of the robot's drivers. The robot is

July 1988

**Optimal Reduced-Order State Estimation
for
Unstable Plants**

by

Dennis S. Bernstein
Harris Corporation
Government Aerospace Systems Division
MS 22/4848
Melbourne, FL 32902

Wassim M. Haddad
Department of Mechanical and
Aerospace Engineering
Florida Institute of Technology
Melbourne, FL 32901

Abstract

The problem of optimal reduced-order steady-state state estimation is considered for the case in which the plant has unstable poles. In contrast to the standard full-order estimation problem involving a single algebraic Riccati equation, the solution to the reduced-order problem involves one modified Riccati equation and one Lyapunov equation coupled by a projection matrix. This projection is completely distinct from the projection obtained in Bernstein and Hyland, 1985, for stable plants.

Supported in part by the Air Force Office of Scientific Research under contract F49620-86-C-0002.

1. Introduction

It has recently been shown that optimal reduced-order, steady-state state estimators can be characterized by means of an algebraic system of equations consisting of one modified Riccati equation and two modified Lyapunov equations coupled by a projection matrix. The solution given in Bernstein and Hyland, 1985, however, was confined to problems in which the plant is asymptotically stable, while in practice it is often necessary to obtain estimators for plants with unstable modes. The purpose of the present paper is to obtain results similar to those of Bernstein and Hyland, 1985, for unstable plants.

Intuitively, it is clear that finite, steady-state state-estimation error for unstable plants is achievable only when the estimator retains, or duplicates in some sense, the unstable modes. Roughly speaking, the solution given in Bernstein and Hyland, 1985, is inapplicable to the unstable problem for the simple reason that the range of the projection matrix may not fully encompass the unstable subspace. Hence, in the present paper we derive a new reduced-order solution which is *constrained* to estimate all of the unstable states. Specifically, for a plant with an unstable subspace of dimension n_u , we characterize the optimal estimator of order n_u which observes all of the unstable states.

As in Bernstein and Hyland, 1985, the solution is given in terms of an oblique projection (denoted in the present paper by μ) which characterizes the optimal estimator gains. Again in contrast to the lone observer Riccati equation of the standard full-order theory, the optimal reduced-order estimator gains for an unstable plant are given by an algebraic system which, in the present case, consists of one modified Riccati equation and one Lyapunov equation coupled by the projection matrix μ .

It is important to stress that the solution derived in the present paper is fundamentally different from the solution obtained in Bernstein and Hyland, 1985, for two reasons. First, the estimator obtained in Bernstein and Hyland, 1985, was characterized by *three* matrix equations (in variables Q , \hat{Q} , and \hat{P}) while the solution obtained herein involves *two* matrix equations (in variables Q and P). And, second, since the projection μ arising in the present paper depends upon P , it is completely distinct from the projection τ appearing in Bernstein and Hyland, 1985, which depends upon \hat{Q} and \hat{P} . Hence the results of the present paper neither generalize, nor are a special case of, the results of Bernstein and Hyland, 1985.

In applying the results of the present paper we note that the solution is applicable to problems in which the unstable subspace also includes additional stable modes. Indeed, the only constraint in applying the theory is that the unstable subspace include all of the unstable poles. To clarify this point (see Sections 2 and 3 for notation) we note that all unstable poles of A must be contained in A_u , but A_u may also contain an arbitrary number of selected stable poles. Thus, the estimator derived in the present paper can be viewed as a subspace-constrained observer-estimator.

Finally, the result given herein is only a partial solution to the reduced-order estimation problem. Specifically, a reduced-order estimator which includes *all* of the unstable modes and optimal combinations of a fixed number of stable modes should involve *both* projections r and μ and *four* matrix equations in variables Q, P, \hat{Q} , and \hat{P} . When the result is specialized to the full-order case we expect the two projections to merge to form the identity and the four matrix equations to collapse to the single observer Riccati equation. A third projection ν due to singular measurement noise and static estimation can also be incorporated (Haddad and Bernstein, 1987, Halevi, 1988). This general solution remains the subject of current research.

After introducing notation, we consider the reduced-order estimation problem for continuous-time plants in Section 2. In Section 3 the corresponding discrete-time problem is considered. For stable plants the reduced-order discrete-time solution was given in Bernstein, Davis, and Hyland, 1986.

Notation and Definitions

Note: All matrices have real entries

| | |
|---|--|
| $\mathbb{R}, \mathbb{R}^{r \times s}, \mathbb{R}^r, \mathbb{E}$ | real numbers, $r \times s$ real matrices, $\mathbb{R}^{r \times 1}$, expected value |
| $I_n, (\)^T, 0_{r \times s}, 0_r$ | $n \times n$ identity matrix, transpose, $r \times s$ zero matrix, $0_r \times r$ |
| $n, \ell, n_e, n_u, n_s, q$ | positive integers |
| x, y, x_e, x_u, x_s, y_e | $n, \ell, n_e, n_u, n_s, q$ -dimensional vectors |
| A, C | $n \times n, \ell \times n$ matrices |
| A_u, A_{us}, A_s | $n_u \times n_u, n_u \times n_s, n_s \times n_s$ matrices |
| C_u, C_s | $\ell \times n_u, \ell \times n_s$ matrices |
| L, L_u, L_s | $q \times n, q \times n_u, q \times n_s$ matrices |
| R | $q \times q$ positive-definite matrix |
| A_e, B_e, C_e, D_e | $n_e \times n_e, n_e \times \ell, q \times n_e, q \times \ell$ matrices |
| t, k | $t \in [0, \infty)$, discrete-time index $1, 2, 3, \dots$ |

| | |
|--------------------------|---|
| \tilde{A} | $A - \begin{bmatrix} I_{n_u} \\ 0_{n_s \times n_u} \end{bmatrix} B_e C$ |
| $w_1(\cdot), w_2(\cdot)$ | n, ℓ -dimensional continuous-time or discrete-time white noise processes |
| V_1 | $n \times n$ nonnegative-definite intensity or covariance of $w_1(\cdot)$ |
| V_2 | $\ell \times \ell$ positive-definite intensity or covariance of $w_2(\cdot)$ |
| V_{12} | $n \times \ell$ cross intensity or cross-covariance of $w_1(\cdot), w_2(\cdot)$ |
| $\tilde{w}(\cdot)$ | $w_1(\cdot) - \begin{bmatrix} I_{n_u} \\ 0_{n_s \times n_u} \end{bmatrix} B_e w_2(\cdot)$ |
| \tilde{V} | $V_1 - V_{12} B_e^T [I_{n_u} \ 0_{n_s \times n_u}] - \begin{bmatrix} I_{n_u} \\ 0_{n_s \times n_u} \end{bmatrix} B_e V_{12}^T + \begin{bmatrix} I_{n_u} \\ 0_{n_s \times n_u} \end{bmatrix} B_e V_2 B_e^T [I_{n_u} \ 0_{n_s \times n_u}]$ |

2. Problem Statement and Main Theorem

Reduced-Order State-Estimation Problem. Given the n th-order observed system

$$\dot{x}(t) = Ax(t) + w_1(t), \quad (2.1)$$

$$y(t) = Cx(t) + w_2(t), \quad (2.2)$$

design an n_s th-order state estimator

$$\dot{x}_e(t) = A_e x_e(t) + B_e y(t), \quad (2.3)$$

$$y_e(t) = C_e x_e(t), \quad (2.4)$$

which minimizes the state-estimation error criterion

$$J(A_e, B_e, C_e) \triangleq \lim_{t \rightarrow \infty} \mathbb{E}[Lx(t) - y_e(t)]^T R [Lx(t) - y_e(t)]. \quad (2.5)$$

In this formulation the plant is partitioned into possibly unstable and stable subsystems. Thus, letting $x(t) = [x_u^T(t), x_s^T(t)]^T$ and $w_1(t) = [w_{1u}^T(t), w_{1s}^T(t)]^T$, (2.1) can be written as

$$\begin{bmatrix} \dot{x}_u(t) \\ \dot{x}_s(t) \end{bmatrix} = \begin{bmatrix} A_u & A_{us} \\ 0_{n_s \times n_u} & A_s \end{bmatrix} \begin{bmatrix} x_u(t) \\ x_s(t) \end{bmatrix} + \begin{bmatrix} w_{1u}(t) \\ w_{1s}(t) \end{bmatrix}, \quad (2.6)$$

where $A_u \in \mathbb{R}^{n_u \times n_u}$ is possibly unstable, $A_s \in \mathbb{R}^{n_s \times n_s}$ is asymptotically stable, and the measurement equation (2.2) becomes

$$y(t) = [C_u \ C_s] \begin{bmatrix} x_u(t) \\ x_s(t) \end{bmatrix} + w_2(t). \quad (2.7)$$

Furthermore, the matrix L , which is partitioned as

$$L = [L_u \ L_s], \quad (2.8)$$

identifies the states or linear combinations of states whose estimates are desired. The dimension n_e of the estimator state x_e is fixed to be equal to the order of the unstable part of the system, i.e., $n_e = n_u$. Thus, the goal of the Reduced-Order State-Estimation Problem is to design an estimator of order n_u which yields quadratically optimal estimates of specified linear combinations of states of the system. As mentioned in Section 1, A_u includes all unstable modes of A as well as an arbitrary number of selected stable modes of A .

Since A_u may contain unstable modes, define the error state $z(t) \triangleq x_u(t) - x_e(t)$ satisfying

$$\dot{z}(t) = (A_u - B_e C_u)x_u(t) - A_e x_e(t) + (A_{us} - B_e C_s)x_s(t) + w_{1u}(t) - B_e w_2(t). \quad (2.9)$$

Note that the explicit dependence of the error states $z(t)$ on the unstable states $x_u(t)$ can be eliminated by constraining

$$A_e = A_u - B_e C_u \quad (2.10)$$

so that (2.9) becomes

$$\dot{z}(t) = (A_u - B_e C_u)z(t) + (A_{us} - B_e C_s)x_s(t) + w_{1u}(t) - B_e w_2(t). \quad (2.11)$$

Similarly, the explicit dependence of the estimation error (2.5) on the unstable states $x_u(t)$ can be eliminated by setting

$$C_e = L_u. \quad (2.12)$$

Now (2.9)–(2.11) yield

$$\dot{\tilde{x}}(t) = \tilde{A}\tilde{x}(t) + \tilde{w}(t), \quad (2.13)$$

where

$$\tilde{x}(t) \triangleq \begin{bmatrix} z(t) \\ x_s(t) \end{bmatrix}, \quad \tilde{A} \triangleq \begin{bmatrix} A_u - B_e C_u & A_{us} - B_e C_s \\ 0_{n_s \times n_u} & A_s \end{bmatrix},$$

and $\tilde{w}(t)$ and its intensity \tilde{V} are given in Section 1.

To guarantee that J is finite, consider the set of asymptotically stable reduced-order estimators

$$S \triangleq \{(A_e, B_e, C_e) : A_e = A_u - B_e C_u \text{ is asymptotically stable}\},$$

so that \tilde{A} is asymptotically stable. Of course, S is nonempty if (A_u, C_u) is detectable. Furthermore, for nondegeneracy we restrict our attention to the set of admissible estimators

$$S^+ \triangleq \{(A_e, B_e, C_e) \in S : (A_e, C_e) \text{ is observable}\},$$

where A_e and C_e are given by (2.10) and (2.12). Also, for arbitrary $Q \in \mathbb{R}^{n \times n}$ define the notation

$$Q_a \triangleq QC^T + V_{12}.$$

Theorem 2.1. Suppose $(A_e, B_e, C_e) \in S^+$ solves the Reduced-Order State-Estimation Problem with constraints (2.10) and (2.12). Then there exist $n \times n$ nonnegative-definite matrices Q, P such that A_e, B_e, C_e are given by

$$A_e = \Phi(A - Q_a V_2^{-1} C) F^T, \quad (2.14)$$

$$B_e = \Phi Q_a V_2^{-1}, \quad (2.15)$$

$$C_e = L F^T, \quad (2.16)$$

and such that Q, P satisfy

$$0 = AQ + QA^T + V_1 - Q_a V_2^{-1} Q_a^T + \mu_{\perp} Q_a V_2^{-1} Q_a^T \mu_{\perp}^T, \quad (2.17)$$

$$0 = (A - \mu Q_a V_2^{-1} C)^T P + P(A - \mu Q_a V_2^{-1} C) + L^T R L, \quad (2.18)$$

where

$$P = \begin{bmatrix} P_u & P_{us} \\ P_{us}^T & P_s \end{bmatrix} \in \mathbb{R}^{(n_u + n_s) \times (n_u + n_s)}, \quad (2.19)$$

$$F \triangleq [I_{n_u} \ 0_{n_u \times n_s}], \quad \Phi \triangleq [I_{n_u} \ P_u^{-1} P_{us}], \quad (2.20)$$

$$\mu \triangleq F^T \Phi = \begin{bmatrix} I_{n_u} & P_u^{-1} P_{us} \\ 0_{n_s \times n_u} & 0_{n_s} \end{bmatrix}, \quad \mu_{\perp} \triangleq I_n - \mu. \quad (2.21)$$

Furthermore, the minimal cost is given by

$$J(A_e, B_e, C_e) = \text{tr } QL^T RL. \quad (2.22)$$

Proof. See Appendix A. \square

Remark 2.1. Note that since $\Phi F^T = I_{n_u}$ the $n \times n$ matrix μ which couples the modified Riccati equation (2.17) and the Lyapunov equation (2.18) is idempotent, i.e., $\mu^2 = \mu$. Note also that $\text{rank } \mu = n_u$. This projection is completely distinct from the projection τ appearing in Bernstein and Hyland, 1985.

Remark 2.2. In the full-order case $n_u = n$, Theorem 2.1 corresponds to the standard steady-state Kalman filter result. To see this, formally set $\Phi = F = \mu = I_n$ and $\mu_\perp = 0_n$ so that (2.18) is superfluous and (2.17) specializes to the standard observer Riccati equation.

Remark 2.3. Note that (2.14) and (2.16) are merely restatements of (2.10) and (2.12). Furthermore, (2.15) implies that $\tilde{A} = A - \mu Q_a V_2^{-1} C$ so that the coefficient of P in (2.18) is asymptotically stable.

3. Discrete-Time Formulation

Discrete-Time Reduced-Order State-Estimation Problem. Given the n th-order observed system

$$x(k+1) = Ax(k) + w_1(k), \quad (3.1)$$

$$y(k) = Cx(k) + w_2(k), \quad (3.2)$$

design an n_e th-order state estimator

$$x_e(k+1) = A_e x_e(k) + B_e y(k), \quad (3.3)$$

$$y_e(k) = C_e x_e(k) + D_e y(k), \quad (3.4)$$

which minimizes the discrete-time state-estimation error criterion

$$\hat{J}(A_e, B_e, C_e, D_e) \triangleq \lim_{k \rightarrow \infty} \mathbb{E}[Lx(k) - y_e(k)]^T R [Lx(k) - y_e(k)]. \quad (3.5)$$

Because of the discrete-time setting it is now possible as in Bernstein, Davis, and Hyland, 1986, to permit a static feedthrough term D_e in the estimator design. The gain D_e represents a *static* least squares estimator in conjunction with the dynamic estimator (A_e, B_e, C_e) .

As in the continuous-time case, the plant is partitioned into stable and possibly unstable subsystems according to (2.6). Furthermore, an error state $z(k) \triangleq x_u(k) - x_e(k)$ is defined, A_e is constrained as in (2.10), and C_e is constrained to be $L_u - D_e C_u$. Thus, the augmented system consisting of the error states $z(k)$ and the stable states $x_s(k)$ becomes

$$\tilde{x}(k+1) = \tilde{A}\tilde{x}(k) + \tilde{w}(k), \quad (3.6)$$

where $\tilde{x}(k) \triangleq [z^T(k), x_s^T(k)]^T$.

To guarantee that J is finite and to obtain closed-form expressions for the estimator gains we restrict our attention to the sets

$$\hat{S} \triangleq \{(A_e, B_e, C_e, D_e) : A_e = A_u - B_e C_u \text{ is asymptotically stable}\},$$

$$\hat{S}^+ \triangleq \{(A_e, B_e, C_e, D_e) \in \hat{S} : (A_e, C_e) \text{ is observable}\}.$$

Also, for arbitrary $Q \in \mathbb{R}^{n \times n}$ define the notation

$$\hat{Q}_a \triangleq AQC^T + V_{12}, \quad \hat{V}_2 \triangleq V_2 + CQC^T.$$

Theorem 3.1. Suppose $(A_e, B_e, C_e, D_e) \in \hat{S}^+$ solves the Discrete-Time Reduced-Order State-Estimation Problem. Then there exist $n \times n$ nonnegative-definite Q, P such that A_e, B_e, C_e, D_e are given by

$$A_e = \Phi(A - \hat{Q}_a \hat{V}_2^{-1} C)F^T, \quad (3.7)$$

$$B_e = \Phi \hat{Q}_a \hat{V}_2^{-1}, \quad (3.8)$$

$$C_e = (L - D_e C)F^T, \quad (3.9)$$

$$D_e = LQC^T \hat{V}_2^{-1}, \quad (3.10)$$

and such that Q, P satisfy

$$Q = AQA^T + V_1 - \hat{Q}_a \hat{V}_2^{-1} \hat{Q}_a^T + \mu_{\perp} \hat{Q}_a \hat{V}_2^{-1} \hat{Q}_a^T \mu_{\perp}^T, \quad (3.11)$$

$$P = (A - \mu \hat{Q}_a \hat{V}_2^{-1} C)^T P (A - \mu \hat{Q}_a \hat{V}_2^{-1} C) + (L - D_e C)^T R (L - D_e C), \quad (3.12)$$

where F, Φ, μ , and μ_{\perp} are defined by (2.19)–(2.21). Furthermore, the minimal cost is given by

$$\hat{J}(A_e, B_e, C_e, D_e) = \text{tr} [(LQL^T - D_e V_2 D_e^T)R]. \quad (3.13)$$

Proof. See Appendix A. \square

Remark 3.1. If a strictly proper estimator is desired, then delete D_e in (3.9), (3.12), and (3.13).

Appendix A: Proof of Theorems 2.1 and 3.1

To analyze (2.13) define the second-moment matrix

$$Q(t) \triangleq \mathbb{E}[\tilde{x}(t)\tilde{x}^T(t)], \quad (\text{A.1})$$

which satisfies

$$\dot{Q}(t) = \tilde{A}Q(t) + Q(t)\tilde{A}^T + \tilde{V}, \quad t \geq 0. \quad (\text{A.2})$$

Since $(A_s, B_s, C_s) \in \mathcal{S}$, \tilde{A} is asymptotically stable and

$$Q \triangleq \lim_{t \rightarrow \infty} \mathbb{E}[\tilde{x}(t)\tilde{x}^T(t)]$$

exists and satisfies

$$0 = \tilde{A}Q + Q\tilde{A}^T + \tilde{V}. \quad (\text{A.3})$$

Next note that (2.5) can be written as

$$J(A_s, B_s, C_s) = \text{tr } QL^T RL. \quad (\text{A.4})$$

To minimize (A.4) over the open set \mathcal{S}^+ subject to the constraint (A.3), form the Lagrangian

$$\mathcal{L}(A_s, B_s, C_s, Q, P, \lambda) \triangleq \text{tr}[\lambda QL^T RL + (\tilde{A}Q + Q\tilde{A}^T + \tilde{V})P], \quad (\text{A.5})$$

where the Lagrange multipliers $\lambda \geq 0$ and $P \in \mathbb{R}^{n \times n}$ are not both zero. Setting $\partial \mathcal{L} / \partial Q = 0$, $\lambda = 0$ implies $P = 0$ since \tilde{A} is asymptotically stable. Hence, without loss of generality set $\lambda = 1$.

Now partition $n \times n$ P into $n_u \times n_u$, $n_u \times n_s$, and $n_s \times n_s$ subblocks as

$$P = \begin{bmatrix} P_u & P_{us} \\ P_{us}^T & P_s \end{bmatrix}. \quad (\text{A.6})$$

Thus the stationarity conditions are given by

$$\frac{\partial \mathcal{L}}{\partial Q} = \tilde{A}^T P + P\tilde{A} + L^T RL = 0, \quad (\text{A.7})$$

$$\frac{\partial \mathcal{L}}{\partial B_s} = P_u B_s V_2 - [P_u \ P_{us}](QC^T + V_{12}) = 0. \quad (\text{A.8})$$

Expanding the $n_u \times n_u$ subblock of (A.7) yields

$$0 = (A_u - B_s C_u)^T P_u + P_u (A_u - B_s C_u) + L_u^T R L_u, \quad (\text{A.9})$$

which, using (2.10) and (2.12), is equivalent to

$$0 = A_e^T P_u + P_u A_e + C_e^T R C_e. \quad (\text{A.10})$$

Thus, since $(A_e, B_e, C_e) \in S^+$, (A_e, C_e) is observable and it follows from (A.10) that P_u is positive definite. Since P_u is thus invertible, define the $n_u \times n$ matrices

$$F \triangleq [I_{n_u} \ 0_{n_u \times n_e}], \quad \Phi \triangleq [I_{n_u} \ P_u^{-1} P_{ue}], \quad (\text{A.11})$$

and the $n \times n$ matrix $\mu \triangleq F^T \Phi$. Note that since $\Phi F^T = I_{n_u}$, μ is idempotent, i.e., $\mu^2 = \mu$.

Next note that (A.8) and (A.11) imply (2.15). Similarly, (2.14) is equivalent to (2.10) with B_e given by (2.15). Finally, (2.16) is a restatement of (2.12). Now, using the expression for B_e , \tilde{A} and \tilde{V} become

$$\tilde{A} = A - \mu Q_a V_2^{-1} C, \quad (\text{A.12})$$

$$\tilde{V} = V_1 - V_{12} V_2^{-1} Q_a^T \mu^T - \mu Q_a V_2^{-1} V_{12}^T + \mu Q_a V_2^{-1} Q_a^T \mu^T. \quad (\text{A.13})$$

Finally, (2.17) and (2.18) follow from (A.3) and (A.7) using (A.12) and (A.13).

For the discrete-time problem define the second-moment matrix

$$Q(k) \triangleq \mathbb{E}[\tilde{x}(k) \tilde{x}^T(k)],$$

which satisfies

$$Q(k+1) = \tilde{A} Q(k) \tilde{A}^T + \tilde{V}. \quad (\text{A.14})$$

Since \tilde{A} is asymptotically stable,

$$Q \triangleq \lim_{k \rightarrow \infty} \mathbb{E}[\tilde{x}(k) \tilde{x}^T(k)]$$

exists and satisfies

$$Q = \tilde{A} Q \tilde{A}^T + \tilde{V}. \quad (\text{A.15})$$

The remainder of the proof follows as above for the continuous-time case.

Acknowledgment. We wish to thank David C. Hyland for several helpful suggestions.

References

- Bernstein, D.S., and Hyland, D.C., 1985, *IEEE Trans. Autom. Contr.*, AC-30, 583.
- Bernstein, D.S., Davis, L.D., and Hyland, D.C., 1986, *AIAA J. Guid. Contr. Dyn.*, 9, 288.
- Haddad, W.M., and Bernstein, D.S., 1987, *IEEE Trans. Autom. Contr.*, AC-32, 1135.
- Halevi, Y., 1988, "The Optimal Reduced-Order Estimator for Systems with Singular Measurement Noise," preprint.

APPENDIX L: Discrete-Time Theory

- 41. D. S. Bernstein, L. D. Davis and D. C. Hyland, "The Optimal Projection Equations for Reduced-Order, Discrete-Time Modelling, Estimation and Control," AIAA J. Guid. Contr. Dyn., Vol. 9, pp. 288-293, 1986.
- 44. D. S. Bernstein, L. D. Davis and S. W. Greeley, "The Optimal Projection Equations for Fixed-Order, Sampled-Data Dynamic Compensation with Computation Delay," IEEE Trans. Autom. Contr., Vol. AC-31, pp. 859-862, 1986.
- 54. W. M. Haddad and D. S. Bernstein, "The Optimal Projection Equations for Discrete-Time Reduced-Order State Estimation for Linear Systems with Multiplicative White Noise," Sys. Contr. Lett., Vol. 8, pp. 381-388, 1987.
- 69. D. S. Bernstein and W. M. Haddad, "Optimal Projection Equations for Discrete-Time Fixed-Order Dynamic Compensation of Linear Systems with Multiplicative White Noise," Int. J. Contr., Vol. 46, pp. 65-73, 1987.
- 92. D. S. Bernstein, "Inequalities for the Trace of Matrix Exponentials," SIAM J. Matrix Anal. Appl., Vol. 9, pp. 156-158, 1988.
- 87. D. S. Bernstein, "Commuting Matrix Exponentials," Problem 88-1, SIAM Review, Vol. 30, p. 123, 1988.
- 128. D. S. Bernstein and C. V. Hollot, "Robust Stability for Sampled-Data Control Systems," Amer. Contr. Conf., Pittsburgh, PA, June 1989.

The Optimal Projection Equations for Reduced-Order, Discrete-Time Modeling, Estimation, and Control

Dennis S. Bernstein,* Lawrence D. Davis,† and David C. Hyland‡
Harris Corporation, Melbourne, Florida

The optimal projection equations derived previously for reduced-order, continuous-time modeling, estimation, and control are developed for the discrete-time case. The design equations are presented in a concise, unified manner to facilitate their accessibility for the development of numerical algorithms for practical applications. As in the continuous-time case, the standard Kalman filter and linear-quadratic-Gaussian results are immediately obtained as special cases of the estimation and control results.

Nomenclature

| | |
|--------------------------------|--|
| A, B, C | $= n \times n, n \times m, \ell \times n$ matrices |
| A_m, B_m, C_m | $= n_m \times n_m, n_m \times m, \ell \times n_m$ matrices |
| A_e, B_e, C_e, D_e | $= n_e \times n_e, n_e \times \ell, p \times n_e, p \times \ell$ matrices |
| A_c, B_c, C_c, D_c | $= n_c \times n_c, n_c \times \ell, m \times n_c, m \times \ell$ matrices |
| E_i | $=$ matrix with unity in the (i, i) position and zeros elsewhere |
| E | $=$ expected value |
| I_r | $= r \times r$ identity matrix |
| k | $=$ discrete-time index $1, 2, 3, \dots$ |
| L | $= p \times n$ matrix |
| $n, m, \ell, n_m, n_e, n_c, p$ | $=$ positive integers, $1 \leq n_m, n_e, n_c \leq n$ |
| R, N, R_2 | $= \ell \times \ell, p \times p, m \times m$ positive-definite matrices |
| R_1 | $= n \times n$ nonnegative-definite matrix |
| R_{12} | $= n \times m$ matrix such that $R_1 - R_{12} R_2^{-1} R_{12}^T$ is nonnegative definite |
| R, R^{*x} | $=$ real numbers, $r \times s$ real matrices |
| \bar{R} | $= \begin{bmatrix} R_1 & R_{12} \\ R_{12}^T & R_2 \end{bmatrix}$ |
| $\text{tr } Z$ | $=$ trace of square matrix Z |
| u, v, y_e | $= m, \ell, p$ -dimensional vectors |
| V | $= m \times m$ positive-definite covariance of w |
| V_1 | $= n \times n$ nonnegative-definite covariance of w_1 |
| V_2 | $= \ell \times \ell$ positive-definite covariance of w_2 |
| V_{12} | $= n \times \ell$ cross-covariance of w_1, w_2 |
| w, w_1, w_2 | $= m, n, \ell$ -dimensional zero-mean discrete-time white noise processes |
| x, x_m, x_e, x_c | $= n, n_m, n_e, n_c$ -dimensional vectors |
| $Z_{(i,j)}$ | $= (i, j)$ element of matrix Z |
| Z^T | $=$ transpose of vector or matrix Z |
| Z^{-T} | $= (Z^T)^{-1}$ or $(Z^{-1})^T$ |
| $\Pi_i(\Psi)$ | $= \Psi E_i \Psi^{-1}$ (unit-rank eigenprojection ⁴⁷) |
| $\rho(Z)$ | $=$ rank of matrix Z |

1. Introduction

IN a recent series of papers¹⁻¹¹ it has been shown that the first-order necessary conditions for quadratically optimal, continuous-time, reduced-order modeling, estimation, and control can be transformed into coupled systems of two, three, and four matrix equations, respectively. This coupling, due to the presence of an oblique projection (idempotent matrix), arises as a rigorous consequence of optimality, hence suggesting the name *optimal projection*. For the estimation and control problems, this formulation provides a direct generalization of classical steady-state Kalman filter and linear-quadratic-Gaussian (LQG) control theory. In the full-order case the projection becomes the identity matrix, the additional two modified Lyapunov equations drop out, and the remaining modified Riccati equations become the usual Riccati equations.

Coupling via the optimal projection supports the view that sequential reduced-order design procedures consisting of either 1) model reduction followed by estimator (controller) design or 2) estimator (controller) design followed by estimator (controller) reduction are generally not optimal. Furthermore, for the control problem the coupled structure of the equations yields the further insight that in the reduced-order case there is no longer separation between the operations of state estimation and state-estimate feedback, i.e., the certainty equivalence principle breaks down.

For practical applications, the optimal projection equations permit the development of alternative numerical algorithms that operate through successive iteration of the optimal projection^{4,6} rather than by gradient search techniques.¹² By recognizing that each local extremal corresponds to n_r possible choices out of n rank-1 eigenprojections of the product of a pair of pseudogramians, it is possible to efficiently identify the global minimum.¹⁰ This idea is philosophically similar to Skelton's component-cost analysis.¹³⁻¹⁵

The purpose of the present paper is to develop the optimal projection equations for reduced-order modeling, estimation, and control in the *discrete-time* case. Since the underlying theory has been discussed previously,¹⁻¹¹ the presentation herein is geared toward a clear and concise statement of the main results to facilitate numerical developments and practical application. For example, by expressing the optimal projection in terms of eigenprojections, a variety of novel algorithms are immediately suggested. For illustrative purposes we apply the results on reduced-order state estimation to a third-order problem to obtain reduced-order estimators and the results on reduced-order dynamic compensation to a tenth-order problem to obtain reduced-order controllers.

Received April 1, 1985; revision received Aug. 26, 1985. Copyright © American Institute of Aeronautics and Astronautics, Inc., 1986. All rights reserved.

*Associate Principal Engineer, Controls Analysis and Synthesis Group, Government Aerospace Systems Division.

†Lead Engineer, Controls Analysis and Synthesis Group, Government Aerospace Systems Division. Member AIAA.

‡Principal Engineer, Controls Analysis and Synthesis Group, Government Aerospace Systems Division. Member AIAA.

Because of the discrete-time setting it is now possible to permit static feedthrough gains in the estimator and controller designs. As previously noted,⁷ nonsingular control weighting and measurement noise in the continuous-time case permit only a purely dynamic (strictly proper) controller. Note that this is precisely the case in continuous-time LQG theory, which always yields a strictly proper feedback controller. The static gains in the discrete-time estimator problem permit simultaneous, unified treatment of nondynamic least-squares estimation along with dynamic (Kalman filter-type) estimation.

The references include a representative sampling of papers on quadratically optimal reduced-order modeling,¹³⁻²⁵ estimation,²⁵⁻³³ and control,³⁴⁻⁴⁶ along with closely related approaches. For emphasis on the discrete-time problem, see Refs. 18, 30, 41, 42, 44, and 45.

II. Problem Statement and Main Results

We now state the reduced-order modeling, estimation, and control problems. The object of the model-reduction problem is to determine a model of reduced state-space dimension whose steady-state response to white noise inputs (or, equivalently, impulse response) best approximates, in a quadratic (least-squares) sense, the response of a given high-order system. In the reduction process the order of the reduced model is fixed and the optimization is performed over the model parameters.

Reduced-Order Modeling Problem

Given the model

$$x(k+1) = Ax(k) + Bw(k) \quad (1)$$

$$y(k) = Cx(k) \quad (2)$$

design a reduced-order model

$$x_m(k+1) = A_m x_m(k) + B_m w(k) \quad (3)$$

$$y_m(k) = C_m x_m(k) \quad (4)$$

which minimizes the model-reduction criterion

$$J_m(A_m, B_m, C_m) \triangleq \lim_{k \rightarrow \infty} E[y_m(k) - y(k)]^T R[y_m(k) - y(k)] \quad (5)$$

The goal of the reduced-order state-estimation problem is to design an estimator of given order which yields quadratically optimal (least squares) estimates of specified linear combinations Lx of states x . In practice, the order of the estimator may be determined by implementation constraints, such as real-time computing capability. Note that the feedthrough term D_e permits the utilization of a static least-squares estimator in conjunction with the dynamic estimator (A_e, B_e, C_e) .

Reduced-Order State-Estimation Problem

Given the observed system

$$x(k+1) = Ax(k) + w_1(k) \quad (6)$$

$$y(k) = Cx(k) + w_2(k) \quad (7)$$

design a reduced-order state estimator

$$x_e(k+1) = A_e x_e(k) + B_e y(k) \quad (8)$$

$$y_e(k) = C_e x_e(k) + D_e y(k) \quad (9)$$

which minimizes the state-estimation criterion

$$J_e(A_e, B_e, C_e, D_e) \triangleq \lim_{k \rightarrow \infty} E[y_e(k) - Lx(k)]^T N[y_e(k) - Lx(k)] \quad (10)$$

For the fixed-order dynamic-compensation problem, a static feedthrough term is included, i.e., the controller may be non-strictly proper.

Reduced-Order Dynamic-Compensation Problem

Given the controlled system

$$x(k+1) = Ax(k) + Bu(k) + w_1(k) \quad (11)$$

$$y(k) = Cx(k) + w_2(k) \quad (12)$$

design a reduced-order dynamic compensator

$$x_c(k+1) = A_c x_c(k) + B_c y(k) \quad (13)$$

$$u(k) = C_c x_c(k) + D_c y(k) \quad (14)$$

which minimizes the dynamic-compensation criterion

$$J_c(A_c, B_c, C_c, D_c) \triangleq \lim_{k \rightarrow \infty} E[x(k)^T R_1 x(k) + 2x(k)^T R_{12} u(k) + u(k)^T R_2 u(k)] \quad (15)$$

To guarantee that J_m , J_e , and J_c are finite and independent of initial conditions, consideration is restricted to the following (open) sets. [A triple $(\hat{A}, \hat{B}, \hat{C})$ is *minimal* if (\hat{A}, \hat{B}) is controllable and (\hat{A}, \hat{C}) is observable.]

$$\mathcal{S}_m \triangleq \{(A_m, B_m, C_m) :$$

$$A_m \text{ is stable and } (A_m, B_m, C_m) \text{ is minimal}\}$$

$$\mathcal{S}_e \triangleq \{(A_e, B_e, C_e, D_e) :$$

$$A_e \text{ is stable and } (A_e, B_e, C_e) \text{ is minimal}\}$$

$$\mathcal{S}_c \triangleq \{(A_c, B_c, C_c, D_c) :$$

$$\begin{bmatrix} A + BD_c C & BC_c \\ B_c C & A_c \end{bmatrix} \text{ is stable and } (A_c, B_c, C_c) \text{ is minimal}\}$$

Let n_r generically denote n_m , n_e , and n_c . The following factorization lemma will be needed for the main results.

Lemma 2.1 Let $\tau \in \mathbb{R}^{n_r \times n_r}$. Then

$$\tau^2 = \tau \quad (16)$$

$$\rho(\tau) = n_r \quad (17)$$

if, and only if, there exist $G, \Gamma \in \mathbb{R}^{n_r \times n_r}$ such that

$$G^T \Gamma = \tau \quad (18)$$

$$\Gamma G^T = I_{n_r} \quad (19)$$

Furthermore, G and Γ are unique to a change of basis in \mathbb{R}^{n_r} .

Proof. Sufficiency is obvious. To prove necessity, first note that due to Eq. (16) the eigenvalues of τ are either 0 or 1. Further, it is easy to see that τ has a diagonal Jordan canonical form. Hence, the result follows from

$$\tau = S \begin{bmatrix} I_{n_r} & 0 \\ 0 & 0 \end{bmatrix} S^{-1} = G^T \Gamma$$

where $G = [\phi^T \ 0] S^T$, $\Gamma = [\phi^{-1} \ 0] S^{-1}$, and $\phi \in \mathbb{R}^{n_r \times n_r}$. ■

For convenience, call G and Γ satisfying Eqs. (18) and (19) a *projective factorization* of τ . Furthermore, for $n \times n$ non-negative-definite matrices (i.e., symmetric matrices with non-negative eigenvalues) \mathcal{Q} and \mathcal{P} , define the set of *contragredi-*

ently diagonalizing transformations

$$\mathcal{D}(\mathcal{Q}, \mathcal{P})$$

$$\triangleq \{ \Psi \in \mathbb{R}^{n \times n} : \Psi^{-1} \mathcal{Q} \Psi^{-T} \text{ and } \Psi^T \mathcal{P} \Psi \text{ are diagonal} \}$$

It follows from Ref. 48, p. 123, Theorem 6.2.5, that $\mathcal{D}(\mathcal{Q}, \mathcal{P})$ is always nonempty. This set does not, however, have a unique element since basis rearrangements and sign transpositions may be incorporated into Ψ . Further nonuniqueness arises if \mathcal{Q} has repeated eigenvalues.

Theorem 2.1. Suppose A is stable and $(A_m, B_m, C_m) \in \mathcal{S}_m$ solves the reduced-order modeling problem. Then there exist $n \times n$ nonnegative-definite matrices \hat{Q} and \hat{P} such that A_m , B_m , and C_m are given by

$$A_m = \Gamma A G^T \quad (20)$$

$$B_m = \Gamma B \quad (21)$$

$$C_m = C G^T \quad (22)$$

and such that \hat{Q} and \hat{P} satisfy

$$\hat{Q} = A \tau \hat{Q} \tau^T A^T + B V B^T \quad (23)$$

$$\hat{P} = A^T \tau^T \hat{P} \tau A + C^T R C \quad (24)$$

where

$$\tau \triangleq \sum_{i=1}^{n_m} \Pi_i(\Psi) \quad (25)$$

for some $\Psi \in \mathcal{D}(\hat{Q}, \hat{P})$ such that $(\Psi^{-1} \hat{Q} \Psi)_{(i,i)} \neq 0$, $i = 1, \dots, n_m$, and some projective factorization G, Γ of τ . Furthermore, the minimal cost is given by

$$J_m(A_m, B_m, C_m) = \text{tr}[(W_c - \tau \hat{Q} \tau^T) C^T R C] \quad (26)$$

where W_c is the unique (nonnegative-definite) solution of

$$W_c = A W_c A^T + B V B^T$$

For convenience in stating the estimator result, define the notation

$$\Sigma_Q \triangleq (A Q C^T + V_{12}) \tilde{V}_2^{-1} (A Q C^T + V_{12})^T$$

$$\Sigma_P \triangleq (L - D_c C)^T N (L - D_c C)$$

$$A_Q \triangleq A - (A Q C^T + V_{12}) \tilde{V}_2^{-1} C$$

$$\tilde{V}_2 \triangleq V_2 + C Q C^T$$

$$\tau_\perp \triangleq I_n - \tau$$

Theorem 2.2. Suppose A is stable and $(A_c, B_c, C_c, D_c) \in \mathcal{S}_c$ solves the reduced-order state-estimation problem. Then there exist $n \times n$ nonnegative-definite matrices Q , \hat{Q} , and \hat{P} such that A_c , B_c , C_c , and D_c are given by

$$A_c = \Gamma [A - (A Q C^T + V_{12}) \tilde{V}_2^{-1} C] G^T \quad (27)$$

$$B_c = \Gamma (A Q C^T + V_{12}) \tilde{V}_2^{-1} \quad (28)$$

$$C_c = (L - D_c C) G^T \quad (29)$$

$$D_c = L Q C^T \tilde{V}_2^{-1} \quad (30)$$

and such that Q , \hat{Q} , and \hat{P} satisfy

$$Q = A Q A^T - (A Q C^T + V_{12}) \tilde{V}_2^{-1} (A Q C^T + V_{12})^T + V_1 + \tau_\perp \hat{Q} \tau_\perp^T \quad (31)$$

$$\hat{Q} = A \tau \hat{Q} \tau^T A^T + \Sigma_Q \quad (32)$$

$$\hat{P} = A^T \tau^T \hat{P} \tau A + \Sigma_P \quad (33)$$

where

$$\tau \triangleq \sum_{i=1}^{n_c} \Pi_i(\Psi) \quad (34)$$

for some $\Psi \in \mathcal{D}(\hat{Q}, \hat{P})$ such that $(\Psi^{-1} \hat{Q} \Psi)_{(i,i)} \neq 0$, $i = 1, \dots, n_c$, and some projective factorization G, Γ of τ . Furthermore, the minimal cost is given by

$$J_c(A_c, B_c, C_c, D_c) = \text{tr}[(L Q L^T - D_c \tilde{V}_2 D_c^T) N] \quad (35)$$

For the control result, define the additional notation

$$\Sigma'_Q \triangleq (A Q C^T + V_{12} + B D_c \tilde{V}_2) \tilde{V}_2^{-1} (A Q C^T + V_{12} + B D_c \tilde{V}_2)^T$$

$$\Sigma'_P \triangleq (B^T P A + R_{12}^T + \tilde{R}_2 D_c C)^T \tilde{R}_2^{-1} (B^T P A + R_{12}^T + \tilde{R}_2 D_c C)$$

$$A_P \triangleq A - B \tilde{R}_2^{-1} (B^T P A + R_{12}^T)$$

$$\tilde{R}_2 \triangleq R_2 + B^T P B$$

$$M \triangleq \begin{bmatrix} I_n \\ D_c C \end{bmatrix}, \quad \hat{M} \triangleq \begin{bmatrix} I_n \\ -\tilde{R}_2^{-1} (B^T P A + R_{12}^T) \end{bmatrix}$$

Theorem 2.3. Suppose $(A_c, B_c, C_c, D_c) \in \mathcal{S}_c$ solves the reduced-order dynamic-compensation problem. Then there exist $n \times n$ nonnegative-definite matrices Q , P , \hat{Q} , and \hat{P} such that A_c , B_c , C_c , and D_c are given by

$$A_c = \Gamma [A - (A Q C^T + V_{12}) \tilde{V}_2^{-1} C - B \tilde{R}_2^{-1} (B^T P A + R_{12}^T) - B D_c] G^T \quad (36)$$

$$B_c = \Gamma [(A Q C^T + V_{12}) \tilde{V}_2^{-1} + B D_c] \quad (37)$$

$$C_c = -[\tilde{R}_2^{-1} (B^T P A + R_{12}^T) + D_c C] G^T \quad (38)$$

$$D_c = -\tilde{R}_2^{-1} (B^T P A Q C^T + R_{12}^T Q C^T + B^T P V_{12}) \tilde{V}_2^{-1} \quad (39)$$

and such that Q , P , \hat{Q} , and \hat{P} satisfy

$$Q = A Q A^T - (A Q C^T + V_{12}) \tilde{V}_2^{-1} (A Q C^T + V_{12})^T + V_1 + \tau_\perp \hat{Q} \tau_\perp^T \quad (40)$$

$$P = A^T P A - (B^T P A + R_{12}^T)^T \tilde{R}_2^{-1} (B^T P A + R_{12}^T) + R_1 + \tau_\perp^T \hat{P} \tau_\perp \quad (41)$$

$$\hat{Q} = A \tau \hat{Q} \tau^T A^T + \Sigma'_Q \quad (42)$$

$$\hat{P} = A^T \tau^T \hat{P} \tau A + \Sigma'_P \quad (43)$$

where

$$\tau \triangleq \sum_{i=1}^{n_c} \Pi_i(\Psi) \quad (44)$$

for some $\Psi \in \mathcal{D}(\hat{Q}, \hat{P})$ such that $(\Psi^{-1} \hat{Q} \Psi)_{(i,i)} \neq 0$, $i = 1, \dots, n_c$, and some projective factorization G, Γ of τ . Further-

more, the minimal cost is given by

$$J(A_c, B_c, C_c, D_c) = \text{tr}[(MQM^T + \hat{M}\tau\hat{Q}\tau^T\hat{M}^T)\hat{R}] \quad (45)$$

Remark 2.1. To specialize the estimation and control results to the strictly proper (no-feedthrough) case, merely ignore Eqs. (30) and (39) and set $D_c = 0$ and $D_e = 0$ wherever they appear.

Remark 2.2. In the full-order cases $n_c = n$ and $n_e = n$ in Theorems 2.2 and 2.3, the projection τ becomes the identity and Eqs. (32), (33), (42), and (43) play no role. In this case G and Γ are also the identity. Specializing further to the purely dynamic case $D_c = 0$, $D_e = 0$ as in the previous remark yields the standard Kalman filter and LQG results.

Remark 2.3. As previously noted,^{7,10,11} the indeterminacy in specifying the projective factorization G, Γ satisfying Eqs. (18) and (19) corresponds to nothing more than an arbitrary choice of internal state-space basis for the design systems (A_n, B_n, C_n) , (A_c, B_c, C_c) , and (A_e, B_e, C_e) .

Remark 2.4. Since \hat{Q} and \hat{P} are balanced by means of the transformation $\Psi \in \mathcal{D}(\hat{Q}, \hat{P})$, it follows that $\Psi^{-1}\hat{Q}\Psi$ is diagonal. Hence, $\hat{Q}\hat{P}$ is semisimple and thus $\Pi_1(\Psi)$ is a rank-1 eigenprojection of $\hat{Q}\hat{P}$. (A semisimple matrix possesses a diagonal Jordan form.^{47,48}) Although the optimal projection τ is characterized in Eqs. (25), (34), and (44) as the sum of rank-1 eigenprojections of $\hat{Q}\hat{P}$, because of the nonuniqueness in $\mathcal{D}(\hat{Q}, \hat{P})$, the theorems do not specify which eigenprojections actually comprise τ . From analytical examples¹⁰ it can be seen that each of the $\binom{n}{n_e}$ possible projections may correspond to a local extremal in the optimization problem.

Remark 2.5. The proofs of Theorems 2.1-2.3 are similar to the continuous-time results and, hence, have been omitted. To help the reader reconstruct the lengthy manipulations, the key details differing from the continuous-time case are pointed out. For the control problem, an $(n + n_c) \times (n + n_c)$ discrete-time algebraic Lyapunov equation is obtained for the steady-state covariance of the closed-loop system. Regarding this equation as a side constraint, the Lagrange multiplier technique is used to compute stationarity conditions that yield explicit expressions for A_c , B_c , C_c , and D_c . The projection arises when these expressions are substituted into the original

augmented Lyapunov equation and its dual. The interesting aspect is that the explicit gain expressions and the definition of the optimal projection arise in the reverse order as compared to the continuous time derivation. Similar remarks apply to the reduced-order modeling and estimation problems.

III. Examples

As an application of Theorem 2.2 on reduced-order state estimation, the stirred-tank example from Ref. 36, pp. 107, 473, and 531, is considered. Ignoring the undisturbed volume state, the remaining states are the incremental tank concentration and variations in the feed concentrations. The problem data are as follows:

$$A = \begin{bmatrix} 0.9048 & 0.06702 & 0.02262 \\ 0 & 0.8825 & 0 \\ 0 & 0 & 0.9048 \end{bmatrix}$$

$$C = [1 \ 0 \ 0], \quad V_2 = 10^{-6}, \quad V_{12} = 0, \quad N = L = I_3$$

$$V_1 = \begin{bmatrix} 5.399 \times 10^{-6} & 8.015 \times 10^{-5} & 8.762 \times 10^{-5} \\ 8.015 \times 10^{-5} & 2.212 \times 10^{-3} & 0 \\ 8.762 \times 10^{-5} & 0 & 7.251 \times 10^{-3} \end{bmatrix}$$

The standard Kalman filter result is

$$A_e = \begin{bmatrix} -0.7959 & 0.06702 & 0.2262 \\ -8.205 & 0.8825 & 0 \\ -10.29 & 0 & 0.9048 \end{bmatrix}$$

$$B_e = \begin{bmatrix} 1.701 \\ 8.205 \\ 10.29 \end{bmatrix}$$

$$C_e = I_3, \quad D_e = 0$$

with performance $J_e = 0.0358515$.

Permitting nonzero feedthrough D_e , yields

$$A_e = \begin{bmatrix} 0.09885 & 0.1061 & -0.0167 \\ -0.6851 & 0.02176 & 0.09015 \\ 0.1503 & 0.2652 & 0.8707 \end{bmatrix}, \quad B_e = \begin{bmatrix} -13.34 \\ -10.08 \\ 2.237 \end{bmatrix}$$

$$C_e = \begin{bmatrix} -0.006261 & -0.007282 & 0.0007445 \\ 0.5716 & -0.0002166 & 0.007202 \\ 0.6822 & -0.006021 & -0.005440 \end{bmatrix},$$

$$D_e = \begin{bmatrix} 0.9068 \\ 9.298 \\ 11.37 \end{bmatrix}$$

where the (improved) performance is $J_e = 0.032401049$.

For the reduced-order results, an algorithm for solving all three equations (31-33) is described briefly. Begin by setting $\tau = I_n$ and solving Eqs. (31-33) for the "full-order" values of Q , \hat{Q} , and \hat{P} . Choose n_e eigenprojections of $\hat{Q}\hat{P}$ in diagonalizing coordinates and iterate the modified Lyapunov equations (32) and (33) until convergence of τ , \hat{Q} , and \hat{P} is obtained. Return to Eq. (31) and solve for Q with $V_1 + \tau_+ \hat{Q} \tau_+^T$ as the new nonhomogeneous term in the Riccati equation. Repeat the above steps until convergence is reached.

In applying this algorithm to the present example, the eigenprojections were chosen for convenience in accordance with the largest eigenvalues of $\hat{Q}\hat{P}$. The results indicate attainment of the global minimum. For the optimal second-order filter, the gains are given by

$$A_e = \begin{bmatrix} 0.09898 & -0.1137 \\ 0.6632 & -0.02285 \end{bmatrix}, \quad B_e = \begin{bmatrix} -13.33 \\ 9.762 \end{bmatrix}$$

$$C_e = \begin{bmatrix} -0.006259 & 0.007796 \\ 0.5716 & -0.0002737 \\ 0.6823 & 0.006121 \end{bmatrix}, \quad D_e = \begin{bmatrix} 0.9068 \\ 9.298 \\ 11.37 \end{bmatrix}$$

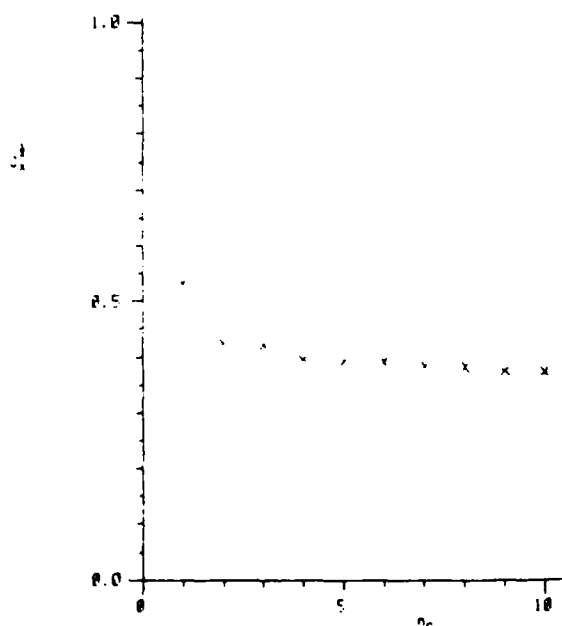


Fig. 1 Root-mean-square performance vs controller order for five-mode beam example.

with $J_c = 0.032401094$, and, for the first-order filter,

$$A_c = 0.1498, \quad B_c = -14.82$$

$$C_c = \begin{bmatrix} -0.001661 \\ 0.5748 \\ 0.6897 \end{bmatrix}, \quad D_c = \begin{bmatrix} 0.9749 \\ 9.489 \\ 11.65 \end{bmatrix}$$

with $J_c = 0.03240418$. Convergence to this accuracy was obtained with 7 iterations of Eqs. (31-33) for the second-order filter and 10 iterations for the first-order filter. Note that the performance degrades only slightly with reduced order, and the static gain term gives the first-order filter better performance than the standard (full-order) Kalman filter.

To illustrate Theorem 2.3 for designing reduced-order dynamic compensators, consider a simply supported beam with two colocated sensor/actuator pairs. Assuming the beam has length 2 and that the sensor/actuator pairs are placed at coordinates $a = 55/172$ and $b = 46/43$, a continuous-time model of the following form is obtained:

$$\dot{x} = \hat{A}x + \hat{B}u + \hat{w}_1 \quad y = \hat{C}x + \hat{w}_2$$

where, retaining the first five modes,

$$\hat{A} = \text{block-diag} \left(\begin{bmatrix} 0 & 1 \\ -\omega_1^2 & -2\zeta\omega_1 \end{bmatrix}, \dots, \begin{bmatrix} 0 & 1 \\ -\omega_5^2 & -2\zeta\omega_5 \end{bmatrix} \right)$$

$$\omega_i = i^2; \quad i = 1, \dots, 5, \quad \zeta = 0.005$$

$$\hat{B}_{(1,1)} = 0.5(1 + (-1)^i) \sin(i\pi b/2), \quad i = 1, \dots, 10$$

$$\hat{B}_{(1,2)} = -0.5(1 + (-1)^i) \sin(i\pi a/2), \quad i = 1, \dots, 10$$

$$\hat{C} = \hat{B}^T$$

The intensities \hat{V}_1 and \hat{V}_2 of \hat{w}_1 and \hat{w}_2 are chosen to be

$$\hat{V}_1 = 0.1I_{10}, \quad \hat{V}_2 = 0.01I_2$$

and it is assumed that \hat{w}_1 and \hat{w}_2 are uncorrelated. For the continuous-time cost

$$J = \lim_{t \rightarrow \infty} E \left[x^T \hat{R}_1 x + 2x^T \hat{R}_{12} u + u^T \hat{R}_2 u \right]$$

set

$$\hat{R}_1 = \text{block-diag} \left(\begin{bmatrix} 1 & 0 \\ 0 & 1/\omega_1 \end{bmatrix}, \dots, \begin{bmatrix} 1 & 0 \\ 0 & 1/\omega_5 \end{bmatrix} \right)$$

$$\hat{R}_{12} = 0, \quad \hat{R}_2 = I_2$$

To convert to the discrete-time problem with discretization interval h , let^{36,49}

$$A = e^{\hat{A}h}, \quad B = \int_0^h e^{\hat{A}t} \hat{B} dt, \quad C = \hat{C}$$

$$V_1 = \int_0^h e^{\hat{A}t} \hat{V}_1 e^{\hat{A}^T t} dt, \quad V_2 = \hat{V}_2$$

$$R_1 = \hat{R}_1, \quad R_{12} = 0, \quad R_2 = \hat{R}_2$$

The design equations (40-43) for the control problem can be solved using exactly the same techniques as in the previous example for the estimation problem. For the strictly proper case ($D_c = 0$), a series of controllers was designed with $n_c = 1, 2, \dots, 10$, where the $n_c = 10$ result is the LQG solution. The

gains for the case $n_c = 4$, for example, are given by

$$A_c = \begin{bmatrix} 0.9317 & 0.1572 & -0.2130 & -0.005038 \\ 0.0137 & 0.6879 & 0.2519 & 0.4085 \\ 0.3330 & -0.0580 & 0.7713 & -0.2602 \\ 0.05980 & -0.3297 & 0.3918 & 0.3005 \end{bmatrix}$$

$$B_c = \begin{bmatrix} -0.4920 & -0.2166 \\ 0.6179 & -0.5959 \\ 0.2253 & -0.02572 \\ 0.07221 & -0.4863 \end{bmatrix}$$

$$C_c = \begin{bmatrix} 0.05864 & -0.3094 & -0.01815 & 0.2409 \\ -0.1301 & 0.1463 & -0.1945 & -0.07192 \end{bmatrix}$$

Figure 1 summarizes the results for each order, where the rms controller performance is given by

$$J_c^i = E \left[\lim_{k \rightarrow \infty} x(k)^T R_1 x(k) \right]^i$$

These results provide a tradeoff study of performance versus controller order that can be used to assess processor requirements.

IV. Conclusion

Optimality conditions have been obtained for the problems of least-squares, reduced-order (i.e., fixed-order), discrete-time modeling, estimation, and control. These conditions comprise systems of two, three, and four matrix equations, respectively, coupled by an oblique projection which determines the optimal system gains. When the order of the estimator or controller is equal to the order of the plant, the oblique projection becomes the identity matrix and the estimation and control results specialize to the standard discrete-time Kalman filter and linear-quadratic-Gaussian results. The design results are applied to two illustrative examples. For a third-order stirred-tank problem, filters of first and second order are obtained, and, for a simply supported Euler beam example with five flexible modes (i.e., 10 states), a series of reduced-order controllers with 1, 2, ..., 9 poles is obtained. The latter results illustrate the tradeoff between control-system performance and controller order.

Acknowledgment

The authors wish to thank Scott W. Greeley for carrying out the computations in Sec. III.

References

- Hyland, D.C., "Optimality Conditions for Fixed-Order Dynamic Compensation of Flexible Spacecraft with Uncertain Parameters," AIAA Paper 82-0312, Jan. 1982.
- Hyland, D.C. and Madiwale, A.N., "Fixed-Order Dynamic Compensation Through Optimal Projection," *Proceedings of the Workshop on Applications of Distributed System Theory to the Control of Large Space Structures*, edited by G. Rodriguez, Jet Propulsion Laboratory, Pasadena, CA, July 1982, pp. 409-427.
- Hyland, D.C., "Mean-Square Optimal Fixed-Order Compensation—Beyond Spillover Suppression," AIAA Paper 82-1403, Aug. 1982.
- Hyland, D.C., "The Optimal Projection Approach to Fixed-Order Compensation: Numerical Methods and Illustrative Results," AIAA Paper 83-0303, Jan. 1983.
- Hyland, D.C. and Bernstein, D.S., "Explicit Optimality Conditions for Fixed-Order Dynamic Compensation," *Proceedings of the 22nd IEEE Conference on Decision and Control*, San Antonio, TX, Dec. 1983, pp. 161-165.
- Hyland, D.C., "Comparison of Various Controller-Reduction Methods: Suboptimal Versus Optimal Projection," *Proceedings of the AIAA Dynamics Specialists Conference*, Palm Springs, CA, May 1984, pp. 381-389.
- Hyland, D.C. and Bernstein, D.S., "The Optimal Projection Equations for Fixed-Order Dynamic Compensation," *IEEE Transactions on Automatic Control*, Vol. AC-29, 1984, pp. 1034-1037.

- ⁸Hyland, D.C. and Bernstein, D.S., "The Optimal Projection Approach to Model Reduction and the Relationship Between the Methods of Wilson and Moore," *Proceedings of the 23rd IEEE Conference on Decision and Control*, Las Vegas, NV, Dec. 1984, pp. 120-126.
- ⁹Bernstein, D.S. and Hyland, D.C., "The Optimal Projection Equations for Reduced-Order State Estimation," *IEEE Transactions on Automatic Control*, Vol. AC-30, 1985, pp. 583-585.
- ¹⁰Hyland, D.C. and Bernstein, D.S., "The Optimal Projection Equations for Model Reduction and the Relationships Among the Methods of Wilson, Skelton and Moore," *IEEE Transactions on Automatic Control*, to appear.
- ¹¹Bernstein, D.S. and Hyland, D.C., "The Optimal Projection Equations for Finite-Dimensional Fixed-Order Dynamic Compensation of Infinite-Dimensional Systems," *SIAM Journal on Control and Optimization*, to appear.
- ¹²Looze, D.P. and Sandell, N. R. Jr., "Gradient Calculations for Linear Quadratic Fixed Control Structure Problems," *IEEE Transactions on Automatic Control*, Vol. AC-25, 1980, pp. 258-285.
- ¹³Skelton, R.E., "Cost Decomposition of Linear Systems with Application to Model Reduction," *International Journal on Control*, Vol. 32, 1980, pp. 1031-1055.
- ¹⁴Skelton, R.E. and Yousuff, A., "Component Cost Analysis of Large Scale Systems," *International Journal on Control*, Vol. 37, 1983, pp. 285-304.
- ¹⁵Yousuff, A. and Skelton, R.E., "Controller Reduction by Component Cost Analysis," *IEEE Transactions on Automatic Control*, Vol. AC-29, 1984, pp. 520-530.
- ¹⁶Mitra, D., "Analytical Results on the Use of Reduced Models in the Control of Linear Dynamical Systems," *Proceedings of IEE*, Vol. 116, 1969, pp. 1439-1444.
- ¹⁷Wilson, D.A., "Optimum Solution of Model-Reduction Problem," *Proceedings of IEE*, Vol. 117, 1970, pp. 1161-1165.
- ¹⁸Appelovich, J.D., "Approximation of Discrete Linear Systems," *International Journal on Control*, Vol. 17, 1973, pp. 565-575.
- ¹⁹Wilson, D.A., "Model Reduction for Multivariable Systems," *International Journal on Control*, Vol. 20, 1974, pp. 57-64.
- ²⁰Hickin, J. and Sinha, N.K., "Applications of Projective Reduction Methods to Estimation and Control," *Journal on Cybernetics*, Vol. 8, 1978, pp. 159-181.
- ²¹Mishra, R.N. and Wilson, D.A., "A New Algorithm for Optimal Reduction of Multivariable Systems," *International Journal on Control*, Vol. 31, 1980, pp. 443-466.
- ²²Kabamba, P.T., "Model Reduction by Euclidean Methods," *Journal of Guidance and Control*, Vol. 3, 1980, pp. 555-562.
- ²³Moore, B.C., "Principal Component Analysis in Linear Systems: Controllability, Observability, and Model Reduction," *IEEE Transactions on Automatic Control*, Vol. AC-26, 1981, pp. 17-32.
- ²⁴Eitelberg, E., "Model Reduction by Minimizing the Weighted Equation Error," *International Journal on Control*, Vol. 34, 1981, pp. 1113-1123.
- ²⁵Sims, C.S., "Reduced-Order Modelling and Filtering," *Control and Dynamic Systems*, Vol. 18, edited by C.T. Leondes, 1982, pp. 55-103.
- ²⁶Simon, K.W. and Stubberud, A.R., "Reduced Order Kalman Filter," *International Journal on Control*, Vol. 10, 1969, pp. 501-509.
- ²⁷Sims, C.S. and Melsa, J.L., "Specific Optimal Estimation," *IEEE Transactions on Automatic Control*, Vol. AC-14, 1969, pp. 183-186.
- ²⁸Asher, R.B., Herring, K.D., and Ryles, J.C., "Bias Variance and Estimation Error in Reduced Order Filters," *Automatica*, Vol. 12, 1976, pp. 589-600.
- ²⁹Galdos, J.I. and Gustafson, D.E., "Information and Distortion in Reduced-Order Filter Design," *IEEE Transactions on Information Theory*, Vol. IT-23, 1977, pp. 183-194.
- ³⁰Fairman, F.W., "Reduced Order State Estimators for Discrete-Time Stochastic Systems," *IEEE Transactions on Automatic Control*, Vol. AC-22, 1977, pp. 673-675.
- ³¹Sims, C.S. and Asher, R.B., "Optimal and Suboptimal Results in Full- and Reduced-Order Linear Filtering," *IEEE Transactions on Automatic Control*, Vol. AC-23, 1978, pp. 469-472.
- ³²Wilson, D.A. and Mishra, R.N., "Design of Low Order Estimators Using Reduced Models," *International Journal on Control*, Vol. 23, 1979, pp. 447-456.
- ³³Dombrovskii, U.V., "Method of Synthesizing Suboptimal Filters of Reduced Order for Digital Linear Dynamic Systems," *Automation and Remote Control*, Vol. 43, 1982, pp. 1483-1489.
- ³⁴Johnson, T.L. and Athans, M., "On the Design of Optimal Constrained Dynamic Compensators for Linear Constant Systems," *IEEE Transactions on Automatic Control*, Vol. AC-15, 1970, pp. 658-660.
- ³⁵Levine, W.S., Johnson, T.L., and Athans, M., "Optimal Limited State Variable Feedback Controllers for Linear Systems," *IEEE Transactions on Automatic Control*, Vol. AC-16, 1971, pp. 785-793.
- ³⁶Kwakernaak, K. and Sivan, R., *Linear Optimal Control Systems*, Wiley-Interscience, New York, 1972.
- ³⁷Rom, D.B. and Sarachik, P.E., "The Design of Optimal Compensators for Linear Constant Systems with Inaccessible States," *IEEE Transactions on Automatic Control*, Vol. AC-18, 1973, pp. 509-512.
- ³⁸Sidar, M. and Kurtaran, B.Z., "Optimal Low-Order Controllers for Linear Stochastic Systems," *International Journal on Control*, Vol. 22, 1975, pp. 377-387.
- ³⁹Mendel, J.J. and Feather, J., "On the Design of Optimal Time-Invariant Compensators for Linear Stochastic Time-Invariant Systems," *IEEE Transactions on Automatic Control*, Vol. AC-20, 1975, pp. 653-657.
- ⁴⁰Basuthakur, S. and Knapp, C.H., "Optimal Constant Controllers for Stochastic Linear Systems," *IEEE Transactions on Automatic Control*, Vol. AC-20, 1975, pp. 664-666.
- ⁴¹Asher, R.B. and Durrett, J.C., "Linear Discrete Stochastic Control with a Reduced-Order Dynamic Compensator," *IEEE Transactions on Automatic Control*, Vol. AC-21, 1976, pp. 626-627.
- ⁴²Naeije, W.J. and Bosgra, O.H., "The Design of Dynamic Compensators for Linear Multivariable Systems," 1977 IFAC, Fredrickton, Canada, pp. 205-212.
- ⁴³Wenk, C.J. and Knapp, C.H., "Parameter Optimization in Linear Systems with Arbitrarily Constrained Controller Structure," *IEEE Transactions on Automatic Control*, Vol. AC-25, 1980, pp. 496-500.
- ⁴⁴O'Reilly, J., "Optimal Low-Order Feedback Controllers for Linear Discrete-Time Systems," *Control and Dynamic Systems*, Vol. 16, edited by C.T. Leondes, Academic Press, New York, 1980, pp. 335-367.
- ⁴⁵Grimble, M.J., "Reduced-Order Optimal Controller for Discrete-Time Stochastic Systems," *Proceedings of IEE*, Vol. 127, 1980, pp. 55-63.
- ⁴⁶Kabamba, P.T. and Longman, R.W., "An Integrated Approach to Reduced-Order Control Theory," *Optimal Control Applications and Methods*, Vol. 4, 1983, pp. 405-415.
- ⁴⁷Kato, T., *Perturbation Theory for Linear Operators*, 2nd Ed., Springer-Verlag, New York, 1976.
- ⁴⁸Rao, C.R. and Mitra, S.K., *Generalized Inverse of Matrices and Its Applications*, John Wiley & Sons, New York, 1971.
- ⁴⁹Van Loan, C.F., "Computing Integrals Involving the Matrix Exponential," *IEEE Transactions on Automatic Control*, Vol. 23, 1978, pp. 395-404.

The Optimal Projection Equations for Fixed-Order Sampled-Data Dynamic Compensation with Computation Delay

DENNIS S. BERNSTEIN, LAWRENCE D. DAVIS, AND
SCOTT W. GREELEY

Abstract—For an LQG-type sampled-data regulator problem which accounts for computational delay and utilizes an averaging A/D device, the equivalent discrete-time problem is shown to be of increased order due to the inclusion of delayed measurement states. The optimal projection equations for reduced-order, discrete-time compensation are applied to the augmented problem to characterize low-order controllers. The design results are illustrated on a tenth-order flexible beam example.

I. INTRODUCTION

Classical sampled-data control theory has been extensively developed [1]–[7] and is widely used in practical applications. Sampled-data design based upon modern optimal control theory has also been developed, although to a considerably lesser extent [8]–[14]. The goals of the present note are twofold. First, for an LQG-type sampled-data regulator problem which explicitly accounts for computational delay, we obtain an equivalent discrete-time problem (Theorem 2.1 and Corollary 2.1). The timing diagram in Fig. 1 illustrates the unavoidable delay in the feedback loop (see Section II for notation).¹ A salient feature of this problem is that rather than replace the continuous-time white noise measurement model by a discrete-time version (which is often done in the literature since continuous-time white noise cannot be sampled), we employ an averaging-type A/D device as in [8, p. 82] [see (2.5)].

The second goal of the note is to present a novel design procedure which is applicable to the equivalent discrete-time problem, and which thus directly accounts for the delay effects. Since the discrete-time model is of augmented order $n + l$ (n = number of plant states, l = number of measurements), it seems natural to seek dynamic feedback of reduced order. To this end, we apply the optimal projection equations for discrete-time dynamic compensation to the equivalent discrete-time problem to characterize optimal controllers of order $n_c \leq n + l$. These equations, which were previously derived in [17] for the continuous-time case, are discussed in [15], [16]. Note that, in practice, the computational delay (and, hence, sample interval) in real-time controller implementation depends directly upon the controller order n_c . For example, by reducing n_c the sample rate can effectively be increased. Thus, the engineering tradeoffs of performance versus controller order and sample interval can be investigated using the results of this note.

This note also includes formulas for integrals of matrix exponentials arising in the sampled-data/discrete-time conversion, along with an algorithm for solving the optimal projection equations. The results are applied to a tenth-order flexible beam example.

II. SAMPLED-DATA PROBLEM AND EQUIVALENT DISCRETE-TIME FORMULATION

The following notation and definitions will be used throughout.

I_r , $0, \dots, 0$, $r \times r$ identity matrix, $r \times s$ zero matrix, $0_{r \times s}$,
 $Z_{(i,j)}$ (i, j) element of matrix Z

Manuscript received October 4, 1985; revised May 12, 1986. This work was supported in part by the Air Force Office of Scientific Research under Contracts AFOSR F49620-84-C-0015, F49620-86-C-0002, and F49620-86-C-0038.

The authors are with the Controls Analysis and Synthesis Group, Government Aerospace Systems Division, Harris Corporation, Melbourne, FL 32902
IEEE Log Number 8609872.

¹ For simplicity, the timing diagram Fig. 1 applies to the case in which the compensator is strictly proper. In the note the results are stated for the more general case in which a direct (static) feedthrough term $D_c \hat{y}(k)$ is included.

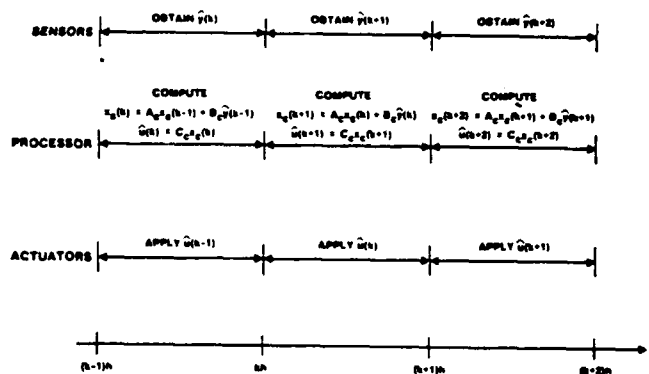


Fig. 1. Timing diagram for sampled-data controller.

| | |
|-------------------------------|--|
| Z^T, Z^{-T} | transpose of vector or matrix Z , $(Z^T)^{-1}$ |
| E_i | matrix with unity in the (i, i) position and zeros elsewhere |
| $E, R, R^{r \times s}$ | expected value, real numbers, $r \times s$ real matrices |
| stable matrix | matrix with eigenvalues in open unit disk |
| nonnegative-semisimple matrix | diagonalizable matrix with nonnegative eigenvalues [18], [19] |
| n, m, l, n_c | positive integers, $1 \leq n_c \leq n + l$ |
| x, u, y, x_c | n, m, l, n_c -dimensional vectors |
| A, B, C | $n \times n, n \times m, l \times n$ matrices |
| A_c, B_c, C_c, D_c | $n_c \times n_c, n_c \times l, m \times n_c, m \times l$ matrices |
| w_1, w_2 | n, l -dimensional zero-mean continuous-time white noise processes |
| V_1 | $n \times n$ nonnegative-definite intensity of w_1 |
| V_2 | $l \times l$ positive-definite intensity of w_2 |
| V_{12} | $n \times l$ cross intensity of w_1, w_2 |
| R_1 | $n \times n$ nonnegative-definite state weighting matrix |
| R_2 | $m \times m$ positive-definite control weighting matrix |
| R_{12} | $n \times m$ cross weighting matrix such that $R_1 - R_{12}R_2^{-1}R_{12}^T$ is nonnegative definite |
| k | discrete-time index, $1, 2, 3, \dots$ |

In the statement of the sampled-data control problem the sample interval h and the controller order n_c are fixed and the optimization is performed over the controller parameters (A_c, B_c, C_c, D_c) . For design tradeoff studies h and n_c can be varied and the problem can be solved for each pair of values of interest.

Fixed-Order, Sampled-Data Dynamic-Compensation Problem

Given the n th-order continuous-time system

$$\dot{x}(t) = Ax(t) + Bu(t) + w_1(t) \quad (2.1)$$

with continuous-time measurements

$$y(t) = Cx(t) + Du(t) + w_2(t) \quad (2.2)$$

design an n_c th-order discrete-time compensator

$$x_c(k+1) = A_c x_c(k) + B_c \hat{y}(k), \quad (2.3)$$

$$\hat{u}(k) = C_c x_c(k) + D_c \hat{y}(k), \quad (2.4)$$

which, with A/D averaged measurements

$$\hat{y}(k) \triangleq \frac{1}{h} \int_{(k-1)h}^{kh} y(t) dt \quad (2.5)$$

and D/A zero-order-hold controls

$$u(t) = \hat{u}(k), \quad t \in [kh, (k+1)h), \quad (2.6)$$

minimizes the performance criterion

$$J(A_c, B_c, C_c, D_c) \triangleq \lim_{t \rightarrow \infty} \frac{1}{t} \int_0^t [x(s)^T R_1 x(s) + 2x(s)^T R_{12} u(s) + u(s)^T R_2 u(s)] ds. \quad (2.7)$$

The main result of this section concerns propagation of the plant and digitized measurements over one time step. For notational convenience, define

$$H(s) \triangleq \int_0^1 e^{-sr} dr.$$

Theorem 2.1: For the fixed-order, sampled-data dynamic-compensation problem, the plant dynamics (2.1), averaged measurements (2.5) and performance criterion (2.7) have the equivalent discrete-time representations

$$x'(k+1) = A' x'(k) + B' \hat{u}(k) + w_1'(k), \quad (2.8)$$

$$\hat{y}(k) = C' x'(k-1) + D' \hat{u}(k-1) + w_2'(k-1), \quad (2.9)$$

$$J(A_c, B_c, C_c, D_c) = \delta + \lim_{k \rightarrow \infty} E[x'(k)^T R_1' x'(k) + 2x'(k)^T R_{12}' \hat{u}(k) + \hat{u}(k)^T R_2' \hat{u}(k)] \quad (2.10)$$

where

$$x'(k) \triangleq x(kh), \quad \delta \triangleq \frac{1}{h} \text{tr} \int_0^h \int_0^s e^{A'V_1} e^{A'T_1} R_1 dr ds,$$

$$A' \triangleq e^{Ah}, \quad B' \triangleq H(h)B, \quad C' \triangleq \frac{1}{h} CH(h), \quad D' \triangleq \frac{1}{h} C \int_0^h H(s) ds B + D;$$

$w_1'(k)$ and $w_2'(k)$ are zero-mean, white noise processes with

$$E \left\{ \begin{bmatrix} w_1'(k) \\ w_2'(k) \end{bmatrix} \begin{bmatrix} w_1'(k) \\ w_2'(k) \end{bmatrix}^T \right\} = \begin{bmatrix} V_1' & V_{12}' \\ V_{12}'^T & V_2' \end{bmatrix},$$

where

$$\begin{aligned} V_1' &\triangleq \int_0^h e^{A'V_1} V_1 e^{A'T_1} ds, \quad V_{12}' \triangleq \frac{1}{h} \int_0^h e^{A'V_1} V_1 H^T(s) ds C^T + \frac{1}{h} H(h) V_{12}, \\ V_2' &\triangleq \frac{1}{h} V_2 + \frac{1}{h^2} C \int_0^h H(s) V_1 H^T(s) ds C^T \\ &\quad + \frac{1}{h^2} C \int_0^h H(s) ds V_{12} + \frac{1}{h^2} V_{12}^T \int_0^h H^T(s) ds C^T; \end{aligned}$$

and

$$\begin{aligned} R_1' &\triangleq \frac{1}{h} \int_0^h e^{A'T_1} R_1 e^{A'V_1} ds, \quad R_{12}' \triangleq \frac{1}{h} \int_0^h e^{A'T_1} R_1 H(s) ds B + \frac{1}{h} H^T(h) R_{12}, \\ R_2' &\triangleq R_2 + \frac{1}{h} B^T \int_0^h H^T(s) R_1 H(s) ds B \\ &\quad + \frac{1}{h} B^T \int_0^h H^T(s) ds R_{12} + \frac{1}{h} R_{12}^T \int_0^h H(s) ds B. \end{aligned}$$

The proof of this theorem is a straightforward calculation involving standard techniques, and hence is omitted. The result is more comprehensive than previous work, however, and includes several results as special cases. For example, the expressions for A' and B' are standard; C' is

given by (10.9), [8, p. 83]; R_1' , R_{12}' , and R_2' are given in [10], [12]; and V_1' , V_{12}' , and V_2' can be found in [8, p. 85]. The expressions for δ and D' appear to be new.

Note that the averaged measurements depend upon delayed samples of the state. By augmenting the discretized state equation (2.8) to include these measurements, it is possible to state the original sampled-data problem as a discrete-time problem.

Corollary 2.1: With the notation

$$\hat{x}(k) \triangleq \begin{bmatrix} x'(k) \\ \hat{y}(k) \end{bmatrix}, \quad \hat{A} \triangleq \begin{bmatrix} A' & 0_{n \times l} \\ C' & 0_l \end{bmatrix}, \quad \hat{B} \triangleq \begin{bmatrix} B' \\ D' \end{bmatrix},$$

$$\hat{C} \triangleq [0_{l \times n} \quad I_l], \quad \hat{w}(k) \triangleq \begin{bmatrix} w_1'(k) \\ w_2'(k) \end{bmatrix}, \quad \hat{V} \triangleq \begin{bmatrix} V_1' & V_{12}' \\ V_{12}'^T & V_2' \end{bmatrix},$$

$$\hat{R}_1 \triangleq \begin{bmatrix} R_1' & 0_{n \times l} \\ 0_{l \times n} & 0_l \end{bmatrix}, \quad \hat{R}_{12} \triangleq \begin{bmatrix} R_{12}' \\ 0_{l \times m} \end{bmatrix}, \quad \hat{R}_2 \triangleq R_2'$$

the fixed-order, sampled-data dynamic-compensation problem is equivalent to the following discrete-time problem. Given the $(n+l)$ -th-order discrete-time system

$$\hat{x}(k+1) = \hat{A} \hat{x}(k) + \hat{B} \hat{u}(k) + \hat{w}(k) \quad (2.11)$$

with discrete-time measurements

$$\hat{y}(k) = \hat{C} \hat{x}(k) \quad (2.12)$$

design an n_c -th-order discrete-time compensator of the form (2.3), (2.4), which minimizes

$$J(A_c, B_c, C_c, D_c) = \delta + \lim_{k \rightarrow \infty} E[\hat{x}(k)^T \hat{R}_1 \hat{x}(k) + 2\hat{x}(k)^T \hat{R}_{12} \hat{u}(k) + \hat{u}(k)^T \hat{R}_2 \hat{u}(k)]. \quad (2.13)$$

Remark 2.1: The equivalent cost (2.13) involves a constant offset δ which serves as a lower bound on the sampled-data performance, i.e., a "discretization floor." Note that

$$\delta = \frac{h}{2} \text{tr} V_1 R_1 + O(h^2).$$

Remark 2.2: Although the measurements $\hat{y}(k)$ are noise free, the singularity is not so serious as singular measurement noise in the continuous-time case where the Kalman filter gains are expressed in terms of the inverse of the measurement noise intensity. In the discrete-time case, rather, it is required that $V + CQC^T$ be invertible, where V is the measurement noise covariance (see [11, p. 530], or [15], [16]).

Remark 2.3: The increase in plant order from n to $n+l$ is due to the computational delay and A/D process. Since discrete-time LQG theory yields a possibly unwieldy $(n+l)$ -th-order controller, we seek "reduced-order" controllers. Note that in this context an n -th-order controller can be regarded as being of reduced order.

Remark 2.4: As pointed out in [10], particular choices of the sample interval h may result in a loss of controllability and observability for the equivalent discrete-time problem. Hence, these properties must be verified before applying control design methods.

III. APPLICATION OF THE OPTIMAL PROJECTION EQUATIONS TO THE EQUIVALENT DISCRETE-TIME PROBLEM

We now apply the optimal projection equations for discrete-time dynamic compensation to the equivalent discrete-time problem. The following easily proved lemma will be needed.

Lemma 3.1: Let $\tau \in R^{(n+l) \times (n+l)}$. Then

$$\tau^2 = \tau, \quad (3.1)$$

$$\rho(\tau) = n_c. \quad (3.2)$$

if and only if there exist $G, \Gamma \in R^{n_c \times (n+1)}$ such that

$$G^T \Gamma = \tau, \quad (3.3)$$

$$\Gamma G^T = I_{n_c}. \quad (3.4)$$

Furthermore, G and Γ are unique to a change of basis in R^{n_c} .

Call G and Γ satisfying (3.3), (3.4) a *projective factorization* of τ . Furthermore, for $n \times n$ nonnegative-definite matrices Q and \mathcal{P} , define the set of *contragrediently diagonalizing transformations*

$$\mathcal{C}(Q, \mathcal{P}) \triangleq \{\Psi \in R^{n \times n}: \Psi^{-1} Q \Psi^{-T} \text{ and } \Psi^T \mathcal{P} \Psi \text{ are diagonal}\}.$$

It follows from [19, Theorem 6.2.5, p. 123] that $\mathcal{C}(Q, \mathcal{P})$ is always nonempty. This set does not, however, have a unique element since basis rearrangements and sign transpositions may be incorporated into Ψ . Further nonuniqueness arises if Q has repeated eigenvalues.

To guarantee that J is finite and independent of initial conditions, we restrict our consideration to the (open) set

$$\mathcal{S} \triangleq \left\{ (A_c, B_c, C_c, D_c) : \begin{bmatrix} \bar{A} + \bar{B} D_c \bar{C} & \bar{B} C_c \\ B_c \bar{C} & A_c \end{bmatrix} \right. \\ \left. \text{is stable and } (A_c, B_c, C_c) \text{ is minimal} \right\}.$$

For the design problem it is required that \mathcal{S} be nonempty, i.e., that the augmented system be stabilizable. We also require the notation

$$\bar{R}_2 \triangleq \bar{R}_2 + \bar{B}^T P \bar{B}, \quad \bar{V}_2 \triangleq \bar{C} Q \bar{C}^T, \quad \tau_{\perp} \triangleq I_{n+1} - \tau,$$

$$A_Q \triangleq \bar{A} - \bar{A} Q \bar{C}^T \bar{V}_2^{-1} \bar{C}, \quad A_P \triangleq \bar{A} - \bar{B} \bar{R}_2^{-1} (\bar{B}^T P \bar{A} + \bar{R}_{12}^T),$$

$$\Sigma_Q \triangleq (\bar{A} Q \bar{C}^T + \bar{B} D_c \bar{V}_2) \bar{V}_2^{-1} (\bar{A} Q \bar{C}^T + \bar{B} D_c \bar{V}_2)^T,$$

$$\Sigma_P \triangleq (\bar{B}^T P \bar{A} + \bar{R}_{12}^T + \bar{R}_2 D_c \bar{C})^T \bar{R}_2^{-1} (\bar{B}^T P \bar{A} + \bar{R}_{12}^T + \bar{R}_2 D_c \bar{C}),$$

$$M \triangleq \begin{bmatrix} I_{n+1} \\ D_c \bar{C} \end{bmatrix}, \quad \bar{M} \triangleq \begin{bmatrix} I_{n+1} \\ -\bar{R}_2^{-1} (\bar{B}^T P \bar{A} + \bar{R}_{12}^T) \end{bmatrix},$$

$$\bar{R} \triangleq \begin{bmatrix} \bar{R}_1 & \bar{R}_{12} \\ \bar{R}_{12}^T & \bar{R}_2 \end{bmatrix}.$$

Theorem 3.1: Suppose $(A_c, B_c, C_c, D_c) \in \mathcal{S}$ solves the fixed-order, sampled-data dynamic-compensation problem. Then there exist $(n+1) \times (n+1)$ nonnegative-definite matrices Q, P, \bar{Q} , and \bar{P} such that A_c, B_c, C_c , and D_c are given by

$$A_c = \Gamma [\bar{A} - \bar{A} Q \bar{C}^T \bar{V}_2^{-1} \bar{C} - \bar{B} \bar{R}_2^{-1} (\bar{B}^T P \bar{A} + \bar{R}_{12}^T) - \bar{B} D_c \bar{C}] G^T, \quad (3.5)$$

$$B_c = \Gamma [\bar{A} Q \bar{C}^T \bar{V}_2^{-1} + \bar{B} D_c], \quad (3.6)$$

$$C_c = -[\bar{R}_2^{-1} (\bar{B}^T P \bar{A} + \bar{R}_{12}^T) + D_c \bar{C}] G^T, \quad (3.7)$$

$$D_c = -\bar{R}_2^{-1} (\bar{B}^T P \bar{A} Q \bar{C}^T + \bar{R}_{12}^T Q \bar{C}^T) \bar{V}_2^{-1}, \quad (3.8)$$

and such that Q, P, \bar{Q} , and \bar{P} satisfy

$$Q = \bar{A} Q \bar{A}^T - \bar{A} Q \bar{C}^T \bar{V}_2^{-1} \bar{C} Q \bar{A}^T + \bar{V} + \tau_{\perp} \bar{Q} \tau_{\perp}^T, \quad (3.9)$$

$$P = \bar{A}^T P \bar{A} - (\bar{B}^T P \bar{A} + \bar{R}_{12}^T)^T \bar{R}_2^{-1} (\bar{B}^T P \bar{A} + \bar{R}_{12}^T) + \bar{R}_1 + \tau_{\perp}^T \bar{P} \tau_{\perp}, \quad (3.10)$$

$$\bar{Q} = \bar{A}_P \tau \bar{Q} \tau^T \bar{A}_P^T + \Sigma_Q, \quad (3.11)$$

$$\bar{P} = \bar{A}_P^T \tau^T \bar{P} \tau \bar{A}_P + \Sigma_P, \quad (3.12)$$

where

$$\tau = \sum_{i=1}^{n_c} \Psi_i E_i \Psi^{-1} \quad (3.13)$$

for some $\Psi \in \mathcal{C}(\bar{Q}, \bar{P})$ such that $(\Psi^{-1} \bar{Q} \Psi^{-T})_{(i,i)} \neq 0, i = 1, \dots, n_c$, and some projective factorization G, Γ of τ . Furthermore, the minimal cost is

given by

$$J(A_c, B_c, C_c, D_c) = \delta + \text{tr} [(MQM^T + \bar{M} \tau \bar{Q} \tau^T \bar{M}^T) \bar{R}]. \quad (3.14)$$

Remark 3.1: Theorem 3.1 can immediately be specialized to the more restrictive problem in which the compensator is strictly proper. This can be done in both the full- and reduced-order cases by ignoring (3.8) and setting $D_c = 0$ wherever it appears. See [15], [16].

IV. NUMERICAL EVALUATION OF INTEGRALS INVOLVING MATRIX EXPONENTIALS

To evaluate the exponential/integral expressions appearing in Theorem 2.1, we utilize the approach of [20]. The idea is to eliminate the need for integration by computing the matrix exponential of appropriate block matrices. Numerical matrix exponentiation is discussed in [21].

Proposition 4.1: Consider the following partitioned matrix exponentials of order $(3n+1) \times (3n+1)$, $(3n+m) \times (3n+m)$, $(2n+m) \times (2n+m)$, and $(3n) \times (3n)$, respectively:

$$\begin{bmatrix} F_1 & F_2 & F_3 & F_4 \\ 0_n & F_5 & F_6 & F_7 \\ 0_n & 0_n & F_8 & F_9 \\ 0_{1 \times n} & 0_{1 \times n} & 0_{1 \times n} & I_1 \end{bmatrix} \triangleq \exp \begin{bmatrix} -A & I_n & 0_n & 0_{n \times 1} \\ 0_n & -A & V_1 & V_{12} \\ 0_n & 0_n & A^T & C^T \\ 0_{1 \times n} & 0_{1 \times n} & 0_{1 \times n} & 0_1 \end{bmatrix} h,$$

$$\begin{bmatrix} F_{10} & F_{11} & F_{12} & F_{13} \\ 0_n & F_{14} & F_{15} & F_{16} \\ 0_n & 0_n & F_{17} & F_{18} \\ 0_{m \times n} & 0_{m \times n} & 0_{m \times n} & I_m \end{bmatrix} \triangleq \exp \begin{bmatrix} -A^T & I_n & 0_n & 0_{n \times m} \\ 0_n & -A^T & R_1 & R_{12} \\ 0_n & 0_n & A & B \\ 0_{m \times n} & 0_{m \times n} & 0_{m \times n} & 0_m \end{bmatrix} h,$$

$$\begin{bmatrix} I_n & F_{19} & F_{20} \\ 0_n & F_{21} & F_{22} \\ 0_{m \times n} & 0_{m \times n} & I_m \end{bmatrix} \triangleq \exp \begin{bmatrix} 0_n & I_n & 0_{n \times m} \\ 0_n & A & B \\ 0_{m \times n} & 0_{m \times n} & 0_m \end{bmatrix} h,$$

$$\begin{bmatrix} F_{23} & F_{24} & F_{25} \\ 0_n & F_{26} & F_{27} \\ 0_n & 0_n & F_{28} \end{bmatrix} \triangleq \exp \begin{bmatrix} -A & I_n & 0_n \\ 0_n & -A & V_1 \\ 0_n & 0_n & A^T \end{bmatrix} h.$$

Then

$$A' = F_{17}, \quad B' = F_{18}, \quad C' = \frac{1}{h} F_{19}^T, \quad D' = \frac{1}{h} C F_{20} + D, \quad \delta = \frac{1}{h} \text{tr } R_1 F_{22}^T F_{25},$$

$$V_1' = F_{14}^T F_6, \quad V_{12}' = \frac{1}{h} F_{14}^T F_7, \quad V_2' = \frac{1}{h} \left(V_2 + \frac{1}{h} C F_{14}^T F_4 + \frac{1}{h} F_{14}^T F_8 C^T \right),$$

$$R_1' = \frac{1}{h} F_{17}^T F_{15}, \quad R_{12}' = \frac{1}{h} F_{17}^T F_{16}, \quad R_2' = R_2 + \frac{1}{h} (B^T F_{17}^T F_{13} + F_{13}^T F_{17} B).$$

V. NUMERICAL SOLUTION OF THE DISCRETE-TIME OPTIMAL PROJECTION EQUATIONS

The following algorithm is proposed for solving (3.9)–(3.12).

Algorithm 5.1:

Step 1: Initialize $k = 0$ and $\tau^{(0)} = I_{n+1}$.

Step 2: With $\tau \triangleq \tau^{(k)}$ solve (3.9)–(3.12) for $Q^{(k)} \triangleq Q, P^{(k)} \triangleq P, \bar{Q}^{(k)} \triangleq \bar{Q}$, and $\bar{P}^{(k)} \triangleq \bar{P}$.

Step 3: If $k \geq 1$ check for convergence: If $\|(Q^{(k)}, P^{(k)}, \bar{Q}^{(k)}, \bar{P}^{(k)}) - (Q^{(k-1)}, P^{(k-1)}, \bar{Q}^{(k-1)}, \bar{P}^{(k-1)})\| > \text{tol}$ then continue; else go to Step 6.

Step 4: Select $\Psi^{(k)} \in \mathcal{C}(\bar{Q}^{(k)}, \bar{P}^{(k)})$ and update $\tau^{(k+1)} = \Sigma_{\tau^{(k)}}, \Psi^{(k)} E_i (\Psi^{(k)})^{-1}$.

Step 5: Increment k and go to Step 2.

Step 6: Evaluate (3.5)–(3.8) with $Q = Q^{(k)}, P = P^{(k)}, \bar{Q} = \bar{Q}^{(k)}, \bar{P} = \bar{P}^{(k)}, G^T \Gamma = \tau^{(k)}, \Gamma G^T = I_{n_c}$.

Remark 5.1: In solving the Riccati equation (3.9), the nonhomogeneous term is taken to be $\bar{V} + \tau_{\perp} \bar{Q} \tau_{\perp}^T$, which is nonnegative definite. Similar remarks apply to (3.10).

TABLE I
COST J FOR BEAM EXAMPLE

| $h \backslash n_c$ | 0 | .1 | .5 |
|--------------------|--------|--------|--------|
| 12 | --- | 1.3715 | 3.0134 |
| 10 | 1.1677 | 1.3723 | 3.0134 |
| 8 | 1.1883 | 1.3823 | 3.0162 |
| 6 | 1.2086 | 1.4421 | 3.0195 |
| 4 | 1.3330 | 1.5752 | 3.0812 |
| 2 | 1.4789 | 2.0425 | 3.3406 |

Open-Loop Cost ($u=0$) is 101.73

Remark 5.2: The critical step of Algorithm 5.1 is the choice of $\Psi^{(k)}$ for constructing the projection $\tau^{(k+1)}$. Since $\Psi^{(k)}$ can include basis rearrangements, the choice of $\Psi^{(k)}$ essentially corresponds to a selection of n_c rank-1 eigenprojections out of $n+1$ possible eigenprojections. This selection is discussed at length in [22] where it is pointed out that the choice of eigenprojections determines which local extremal will be reached by the algorithm. Component-cost methods have thus been utilized as a promising selection criteria. Because of the eigenprojection structure of the necessary conditions, Algorithm 5.1 is fundamentally different from gradient search methods.

VI. ILLUSTRATIVE EXAMPLE: CONTROL OF A FLEXIBLE BEAM

Consider a simply supported beam of length two with two colocated sensor/actuator pairs placed at coordinates $a_1 = 55/172$ and $a_2 = 45/43$. Define

$$A = \text{block-diag}_{i=1, \dots, 5} \begin{bmatrix} 0 & 1 \\ -\omega_i^2 & -2\zeta_i\omega_i \end{bmatrix}, \quad \omega_i = i^2, \quad i = 1, \dots, 5, \quad \zeta = 0.005,$$

$$B_{(i,j)} = 0.5(-1)^{i-1}(1+(-1)^j) \sin(i\pi a_j/4),$$

$$i = 1, \dots, 10, \quad j = 1, 2, \quad C = B^T,$$

$$V_1 = \text{block-diag}_{i=1, \dots, 5} \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}, \quad V_{12} = 0, \quad V_2 = 0.1I_2,$$

$$R_1 = \text{block-diag}_{i=1, \dots, 5} \begin{bmatrix} 1 & 0 \\ 0 & \omega_i^{-2} \end{bmatrix}, \quad R_{12} = 0, \quad R_2 = 0.1I_2.$$

For $n_c = 10, 8, 6, 4, 2$ continuous-time controllers were designed using the results of [17] and, for $n_c = 12, 10, 8, 6, 4, 2$ and $h = 0.1, 0.5$, strictly proper ($D_c = 0$) discrete-time controllers were obtained from Theorem 3.1. The results are summarized in Table I.

VII. DIRECTIONS FOR FURTHER DEVELOPMENTS

The following extensions and related developments immediately suggest themselves: reduced-order, discrete-time modeling/estimation of

continuous-time systems [22], [23]; robust sampled-data control of uncertain systems with multiplicative noise [24]–[27]; multirate sampling [28], [29]; alternative A/D and D/A devices and asynchronous sampling/control update; infinite-dimensional systems [30], [31].

REFERENCES

- [1] J. R. Ragazzini and G. F. Franklin, *Sampled-Data Control Systems*. New York: McGraw-Hill, 1958.
- [2] E. I. Jury, *Sampled-Data Control Systems*. New York: Wiley, 1958.
- [3] J. T. Tou, *Digital and Sampled-Data Control Systems*. New York: McGraw-Hill, 1959.
- [4] A. J. Munroe, *Digital Processes for Sampled Data Systems*. New York: Wiley, 1962.
- [5] B. C. Kuo, *Analysis and Synthesis of Sampled-Data Control Systems*. Englewood Cliffs, NJ: Prentice-Hall, 1963.
- [6] D. P. Lindorff, *Theory of Sampled-Data Control Systems*. New York: Wiley, 1964.
- [7] G. F. Franklin and J. D. Powell, *Digital Control of Dynamic Systems*. Reading, MA: Addison-Wesley, 1980.
- [8] K. J. Astrom, *Introduction to Stochastic Control Theory*. New York: Academic, 1970.
- [9] A. H. Levis, M. Athans, and R. Schlueter, "On the behavior of optimal linear sampled-data regulators," *Int. J. Contr.*, vol. 13, pp. 343–361, 1971.
- [10] P. Dorato and A. Levis, "Optimal linear regulators: The discrete case," *IEEE Trans. Automat. Contr.*, vol. AC-16, pp. 613–620, 1971.
- [11] K. Kwakernaak and R. Sivan, *Linear Optimal Control Systems*. New York: Wiley, 1972.
- [12] N. Halyo and A. K. Caglayan, "A separation theorem for the stochastic sampled-data LQG problem," *Int. J. Contr.*, vol. 23, pp. 237–244, 1976.
- [13] E. S. Armstrong and A. K. Caglayan, "An algorithm for the weighting matrices in the sampled-data optimal linear regulator problem," NASA Langley, Hampton, VA, Rep. NASA TN D-8372, 1976.
- [14] T. Mita, "Optimal digital feedback control systems counting computation time of control laws," *IEEE Trans. Automat. Contr.*, vol. AC-30, pp. 542–548, 1985.
- [15] D. S. Bernstein, L. D. Davis, S. W. Greeley, and D. C. Hyland, "The optimal projection equations for reduced-order, discrete-time modelling, estimation and control," *Proc. 24th IEEE Conf. Dec. Contr.*, Fort Lauderdale, FL, Dec. 1985, pp. 573–578.
- [16] D. S. Bernstein, L. D. Davis, and D. C. Hyland, "The optimal projection equations for reduced-order, discrete-time modelling, estimation and control," *J. Guid. Contr. Dynam.*, vol. 9, pp. 288–293, 1986.
- [17] D. C. Hyland and D. S. Bernstein, "The optimal projection equations for fixed-order dynamic compensation," *IEEE Trans. Automat. Contr.*, vol. AC-29, pp. 1034–1037, 1984.
- [18] T. Kato, *Perturbation Theory for Linear Operators*. 2nd ed. New York: Springer-Verlag, 1976.
- [19] C. R. Rao and S. K. Mitra, *Generalized Inverse of Matrices and Its Applications*. New York: Wiley, 1971.
- [20] C. F. Van Loan, "Computing integrals involving the matrix exponential," *IEEE Trans. Automat. Contr.*, vol. AC-23, pp. 395–404, 1978.
- [21] C. Moler and C. Van Loan, "Nineteen dubious ways to compute the exponential of a matrix," *SIAM Rev.*, vol. 20, pp. 801–836, 1978.
- [22] D. C. Hyland and D. S. Bernstein, "The optimal projection equations for model reduction and the relationships among the methods of Wilson, Skelton, and Moore," *IEEE Trans. Automat. Contr.*, vol. AC-30, pp. 1201–1211, 1985.
- [23] ———, "The optimal projection equations for reduced-order state estimation," *IEEE Trans. Automat. Contr.*, vol. AC-30, pp. 583–585, 1985.
- [24] D. S. Bernstein and D. C. Hyland, "Optimal projection/maximum entropy stochastic modelling and reduced-order design synthesis," in *Proc. IFAC Workshop on Model Error Concepts and Compensation*, Boston, MA, June 1985.
- [25] ———, "The optimal projection/maximum entropy approach to designing low-order, robust controllers for flexible structures," in *Proc. 24th IEEE Conf. Decision Contr.*, Fort Lauderdale, FL, Dec. 1985, pp. 745–752.
- [26] ———, "The optimal projection equations for reduced-order modelling, estimation and control of linear systems with Stratonovich multiplicative white noise," submitted for publication.
- [27] A. R. Tiedemann and W. L. DeKoning, "The equivalent discrete-time optimal control problem for continuous-time systems with stochastic parameters," *Int. J. Contr.*, vol. 70, pp. 449–466, 1984.
- [28] D. P. Glasson, "A new technique for multirate digital control design and sample rate selection," *J. Guid. Contr. Dynam.*, vol. 5, pp. 379–382, 1982.
- [29] M. Araki and K. Yamamoto, "Multivariable multirate sampled-data systems. State-space description, transfer characteristics, and Nyquist criterion," *IEEE Trans. Automat. Contr.*, vol. AC-31, pp. 145–154, 1986.
- [30] M. J. Balas, "The structure of discrete-time finite-dimensional control of distributed parameter systems," *J. Math. Anal. Appl.*, to be published.
- [31] D. S. Bernstein and D. C. Hyland, "The optimal projection equations for finite-dimensional fixed-order dynamic compensation of infinite-dimensional systems," *SIAM J. Contr. Optimiz.*, vol. 24, pp. 122–151, 1986.

The optimal projection equations for reduced-order, discrete-time state estimation for linear systems with multiplicative white noise

Wassim M. HADDAD

Mechanical Engineering Department, Florida Institute of Technology, Melbourne, FL 32901, USA

Dennis S. BERNSTEIN *

Harris Corporation, Government Aerospace Systems Division, MS 22/4842, Melbourne, FL 32902, USA

Received 2 July 1986

Abstract: The optimal projection equations obtained in [2,3] for reduced-order, discrete-time state estimation are generalized to include the effects of state- and measurement-dependent noise to provide a model of parameter uncertainty. In contrast to the single matrix Riccati equation arising in the full-order (Kalman filter) case, the optimal steady-state reduced-order discrete-time estimator is characterized by three matrix equations (one modified Riccati equation and two modified Lyapunov equations) coupled by both an oblique projection and stochastic effects.

Keywords: Reduced-order Kalman filter, Robust estimation.

1. Introduction

In a recent series of papers [1–3] it has been shown that the first-order necessary conditions for optimal continuous and discrete-time reduced-order state-estimation can be transformed into coupled systems of three matrix equations (one modified Riccati equation and two modified Lyapunov equations). The coupling is due to the presence of an oblique projection (idempotent matrix) which arises as a rigorous consequence of the stationarity conditions. This formulation provides a direct generalization of the classical steady-state Kalman filter theory. Specifically, in the full-order case, the projection becomes the identity matrix, the additional two modified Lyapunov equations drop out, and the remaining modified Riccati equation reduces to the standard observer Riccati equation for the Kalman filter expression. Related results in reduced-order estimator design can be found in [4–17].

An additional extension of classical state estimation involves the inclusion of state- and measurement-dependent disturbances [18–24]. One motivation for such a model is to design estimators which are desensitized, i.e., robustified, to *actual* parameter variations [25–31]. For the continuous-time control problem this has been justified in [32–38].

As shown in [36] for the continuous-time case, applying the optimal projection approach to the multiplicative white noise model yields an extended formulation of the optimality conditions for reduced-order state estimation. Specifically, the system of three matrix equations characterizing the optimal estimator are now coupled by both the oblique projection and stochastic effects.

* Supported in part by the Air Force Office of Scientific Research under contract F49620-86-C-0002.

The purpose of the present paper is to provide a self-contained derivation of the optimality conditions for reduced-order state estimation in the presence of both state- and measurement-dependent white noise in the discrete-time case. The goal of the development is to present the optimality conditions in a clear, concise manner to facilitate the development of numerical algorithms for practical application.

2. Notation and definitions

| | |
|---|--|
| $\mathbf{R}, \mathbf{R}^{r \times s}, \mathbf{R}^r, \mathbf{E}$ | real numbers, $r \times s$ real matrices, $\mathbf{R}^{r \times 1}$, expectation. |
| $I_n, (\cdot)^T, \otimes$ | $n \times n$ identity, transpose, Kronecker product [39]. |
| τ_1 | $I_n - \tau, \tau \in \mathbf{R}^{n \times n}$. |
| n, m, l, n_e, q | positive integers, $1 \leq n_e \leq n$. |
| x, x_e | n, n_e -dimensional vectors. |
| y, y_e | l, q -dimensional vectors. |
| A, A_i, C, C_i | $n \times n$ matrices, $l \times n$ matrices, $i = 1, \dots, p$. |
| A_e, B_e, C_e, D_e | $n_e \times n_e, n_e \times l, q \times n_e, q \times l$ matrices. |
| k | discrete-time index $1, 2, 3, \dots$. |
| $v_i(k)$ | unit variance white noise, $i = 1, \dots, p$. |
| $w_1(k), w_2(k)$ | n -dimensional, l -dimensional white noise processes. |
| V_1 | $n \times n$ nonnegative-definite covariance of $w_1(k)$. |
| V_2 | $l \times l$ positive-definite covariance of $w_2(k)$. |
| V_{12} | $n \times l$ cross-covariance of $w_1(k), w_2(k)$. |
| R | $q \times q$ positive-definite matrix. |
| L | $q \times n$ matrix. |

$$\tilde{A} = \begin{bmatrix} A & 0 \\ B_e C & A_e \end{bmatrix}, \quad \tilde{A}_i = \begin{bmatrix} A_i & 0 \\ B_e C_i & 0 \end{bmatrix}, \quad i = 1, \dots, p,$$

$$\tilde{w}(k) = \begin{bmatrix} w_1(k) \\ B_e w_2(k) \end{bmatrix}, \quad \tilde{V} = \begin{bmatrix} V_1 & V_{12} B_e^T \\ B_e V_{12}^T & B_e V_2 B_e^T \end{bmatrix},$$

$$\tilde{R} = \begin{bmatrix} L^T R L - L^T R D_e C - C^T D_e^T R L + C^T D_e^T R D_e C + \sum_{i=1}^p C_i^T D_e^T R D_e C_i & -L^T R C_e + C^T D_e^T R C_e \\ -C_e^T R L + C_e^T R D_e C & C_e^T R C_e \end{bmatrix}.$$

| | |
|----------------------------------|--|
| $Z_{(i,j)}$ | (i, j) element of matrix Z . |
| $\rho(Z)$ | rank of matrix Z . |
| $\text{tr } Z$ | trace of a square matrix Z . |
| E_i | square matrix with unity in the (i, i) position and zeros elsewhere. |
| $\pi_i(\psi)$ | $\psi E_i \psi^{-1}$ (unit-rank eigenprojection). |
| $\mathcal{N}(Z), \mathcal{R}(Z)$ | null space, range of matrix Z . |

An asymptotically stable matrix is a matrix with eigenvalues in the open unit disk; a nonnegative-definite matrix is a symmetric matrix with nonnegative eigenvalues; and a positive-definite matrix is a symmetric matrix with positive eigenvalues.

For arbitrary $n \times n, Q, \hat{Q}, \tau$, define

$$V_{2s} \triangleq V_2 + C Q C^T + \sum_{i=1}^p C_i (Q + \tau \hat{Q} \tau^T) C_i^T, \quad Q_s \triangleq A Q C^T + V_{12} + \sum_{i=1}^p A_i (Q + \tau \hat{Q} \tau^T) C_i^T,$$

$$\hat{V}_{2s} \triangleq V_2 + C Q C^T + \sum_{i=1}^p C_i (Q + \hat{Q}) C_i^T, \quad \hat{Q}_s \triangleq A Q C^T + V_{12} + \sum_{i=1}^p A_i (Q + \hat{Q}) C_i^T.$$

3. Problem statement and main theorem

Reduced-Order State-Estimation Problem. Given the n -th-order observed system

$$x(k+1) = \left(A + \sum_{i=1}^p v_i(k) A_i \right) x(k) + w_1(k), \quad (3.1)$$

$$y(k) = \left(C + \sum_{i=1}^p v_i(k) C_i \right) x(k) + w_2(k), \quad (3.2)$$

design an n_e -th reduced-order state estimator

$$x_e(k+1) = A_e x_e(k) + B_e y(k), \quad (3.3)$$

$$y_e(k) = C_e x_e(k) + D_e y(k), \quad (3.4)$$

which minimizes the state-estimation error criterion

$$J(A_e, B_e, C_e, D_e) \triangleq \lim_{k \rightarrow \infty} E[Lx(k) - y_e(k)]^T R[Lx(k) - y_e(k)]. \quad (3.5)$$

In this formulation the matrix L identifies the states, or linear combinations of states, whose estimates are desired. The order n_e of the estimator state x_e is determined by implementation constraints, i.e., by the computing capability available for realizing (3.3), (3.4) in real time. Note that the feedthrough term D_e permits the utilization of a *static* least squares estimator in conjunction with the dynamic estimator (A_e, B_e, C_e) . Thus, the goal of the Reduced-Order State-Estimation Problem is to design an estimator of given order that yields quadratically optimal (least squares) estimates of specified linear combinations of states.

To guarantee that J is finite assume that A is asymptotically stable and consider the set of asymptotically stable reduced-order (i.e., fixed-order) estimators

$$\mathcal{S} \triangleq \{(A_e, B_e, C_e, D_e) : A_e \text{ is asymptotically stable}\}.$$

Since the value of J is independent of the internal realization of the transfer function corresponding to (3.3) and (3.4), without loss of generality we further restrict our attention to the set of admissible estimators

$$\mathcal{S}^+ \triangleq \{(A_e, B_e, C_e, D_e) \in \mathcal{S} : (A_e, B_e) \text{ is controllable and } (A_e, C_e) \text{ is observable}\}.$$

The following factorization lemma is needed for the statement of the main result.

Lemma 3.1. Let $\tau \in \mathbb{R}^{n \times n}$. Then

$$\tau^{2^*} = \tau, \quad (3.6)$$

$$\rho(\tau) = n_e, \quad (3.7)$$

if and only if there exist $G, \Gamma \in \mathbb{R}^{n_e \times n}$ such that

$$G^T \Gamma = \tau, \quad (3.8)$$

$$\Gamma G^T = I_{n_e}. \quad (3.9)$$

Furthermore, G and Γ are unique to a change of basis in \mathbb{R}^{n_e} .

Proof. See [3]. \square

For convenience, call G and Γ satisfying (3.8) and (3.9) a *projective factorization* of τ . Furthermore, for $n \times n$ nonnegative-definite matrices \hat{Q} and \hat{P} , define the set of *contragrediently diagonalizing* transformations

$$\mathcal{D}(\hat{Q}, \hat{P}) \triangleq \{\psi \in \mathbb{R}^{n \times n}: \psi^{-1}\hat{Q}\psi^{-T} \text{ and } \psi^T\hat{P}\psi \text{ are diagonal}\}.$$

It follows from Theorem 6.2.5, p. 123 of [40], that $\mathcal{D}(\hat{Q}, \hat{P})$ is always nonempty. This set does not, however, have a unique element since basis rearrangements and sign transpositions may be incorporated into ψ . Further nonuniqueness arises if $\hat{Q}\hat{P}$ has repeated eigenvalues.

Theorem 3.1. Suppose A is asymptotically stable and $(A_e, B_e, C_e, D_e) \in \mathcal{S}^+$ solves the Optimal Reduced-Order State-Estimation Problem. Then there exist $n \times n$ nonnegative-definite matrices Q , \hat{Q} and \hat{P} such that A_e , B_e , C_e and D_e are given by

$$A_e = \Gamma(A - Q_s V_{2s}^{-1} C) G^T, \quad (3.10)$$

$$B_e = \Gamma Q_s V_{2s}^{-1}, \quad (3.11)$$

$$C_e = (L - D_e C) G^T, \quad (3.12)$$

$$D_e = L Q C^T V_{2s}^{-1}, \quad (3.13)$$

and such that Q , \hat{Q} and \hat{P} satisfy

$$Q = A Q A^T + \sum_{i=1}^p A_i (Q + \tau \hat{Q} \tau^T) A_i^T + V_1 - Q_s V_{2s}^{-1} Q_s^T + \tau_1 \hat{Q} \tau_1^T, \quad (3.14)$$

$$\hat{Q} = A \tau \hat{Q} \tau^T A^T + Q_s V_{2s}^{-1} Q_s^T, \quad (3.15)$$

$$\hat{P} = (A - Q_s V_{2s}^{-1} C)^T \tau^T \hat{P} \tau (A - Q_s V_{2s}^{-1} C) + (L - D_e C)^T R (L - D_e C), \quad (3.16)$$

where

$$\tau \triangleq \sum_{i=1}^{n_e} \pi_i(\psi) \quad (3.17)$$

for some $\psi \in \mathcal{D}(\hat{Q}, \hat{P})$ such that $(\psi^{-1}\hat{Q}\hat{P}\psi)_{(i,i)} \neq 0$, $i = 1, \dots, n_e$, and some projective factorization G , Γ of τ . Furthermore, the minimal cost is given by

$$J(A_e, B_e, C_e, D_e) = \text{tr}[(L Q L^T - D_e V_{2s} D_e^T) R]. \quad (3.18)$$

Remark 3.1. It is useful to note that (3.10) can be replaced by

$$A_e = \Gamma A G^T - B_e C G^T. \quad (3.10)'$$

Remark 3.2. Because of (3.9) the $n \times n$ matrix τ which couples the three equations (3.14)–(3.16) is idempotent, i.e., $\tau^2 = \tau$. In general, this ‘optimal projection’ is an oblique projection (as opposed to an orthogonal projection) since it is not necessarily symmetric. It should be stressed that the form of the optimal reduced-order estimator (3.10)–(3.13) is a direct consequence of optimality and not the result of an a priori assumption on the structure of the reduced-order estimator.

Remark 3.3. To specialize the result to the strictly proper (no feedthrough) case, merely ignore (3.13) and set $D_e = 0$ wherever it appears.

Remark 3.4. Replacing x_e by Sx_e , where S is invertible, yields the 'equivalent' estimator $(SA_e S^{-1}, SB_e, C_e S^{-1}, D_e)$ with $J(SA_e S^{-1}, SB_e, C_e S^{-1}, D_e) = J(A_e, B_e, C_e, D_e)$. Note that transformation of the estimator state basis corresponds to the alternative factorization $\tau = (S^{-1}G)^T(S\Gamma)$.

Remark 3.5. Note that for the optimal values of A_e and B_e the estimator dynamics (3.3) assume the usual observer form

$$x_e(k+1) = \Gamma A G^T x_e + \Gamma Q_s V_{2s}^{-1} (y - C G^T x_e). \quad (3.19)$$

By introducing the quasi-full-state estimate $\hat{x} \triangleq G^T x_e \in \mathbb{R}^n$ so that $\tau \hat{x} = \hat{x}$ and $x_e = \Gamma \hat{x} \in \mathbb{R}^{n_e}$, (3.19) can be written as

$$\hat{x}(k+1) = \tau A \tau \hat{x} + \tau Q_s V_{2s}^{-1} (y - C \hat{x}). \quad (3.20)$$

Although the implemented estimator (3.19) has the state $x_e \in \mathbb{R}^{n_e}$, (3.19) can be viewed as a quasi-full-order estimator whose geometric structure is entirely dictated by the projection τ and the stochastic effects. Specifically, error inputs $Q_s V_{2s}^{-1} (y - C \hat{x})$ are annihilated unless they are contained in $[\mathcal{N}(\tau)]^\perp = \mathcal{R}(\tau^T)$. Hence, the observation subspace of the estimator is precisely $\mathcal{R}(\tau^T)$.

Specializing Theorem 3.1 to the noise-free case $A_i = 0$, $C_i = 0$, $i = 1, \dots, p$, yields Theorem 2.2 of [2,3]. Alternatively, specializing Theorem 3.1 to the full-order case $n_e = n$ reveals that the Lyapunov equation for \hat{P} is superfluous. In this case it follows from Remark 3.4 that $G = \Gamma = I_n$ without loss of generality.

Corollary 3.1. Assume $n_e = n$, A is asymptotically stable and $(A_e, B_e, C_e, D_e) \in \mathcal{S}^+$ solves the Optimal Full-Order State-Estimation Problem. Then there exist $n \times n$ nonnegative-definite matrices Q and \hat{Q} such that A_e , B_e , C_e and D_e are given by

$$A_e = A - \hat{Q}_s \hat{V}_{2s}^{-1} C, \quad (3.21)$$

$$B_e = \hat{Q}_s \hat{V}_{2s}^{-1}, \quad (3.22)$$

$$C_e = L - D_e C, \quad (3.23)$$

$$D_e = L Q C^T \hat{V}_{2s}^{-1}, \quad (3.24)$$

and such that Q and \hat{Q} satisfy

$$Q = A Q A^T + \sum_{i=1}^p A_i (Q + \hat{Q}) A_i^T + V_1 - \hat{Q}_s \hat{V}_{2s}^{-1} \hat{Q}_s^T, \quad (3.25)$$

$$\hat{Q} = A \hat{Q} A^T + \hat{Q}_s \hat{V}_{2s}^{-1} \hat{Q}_s^T. \quad (3.26)$$

Furthermore, the minimal cost is given by

$$J(A_e, B_e, C_e, D_e) = \text{tr}[(L Q L^T - D_e \hat{V}_{2s} D_e^T) R]. \quad (3.27)$$

Remark 3.6. To recover the standard Kalman filter result from Corollary 3.1 set $A_i = 0$, $C_i = 0$, $i = 1, \dots, p$, so that (3.25) and (3.26) are decoupled and (3.26) is superfluous. Since the standard Kalman filter is strictly proper, set $D_e = 0$ as in Remark 3.3.

4. Proof of the main theorem

Using the notation of Section 2 the augmented system (3.1) and (3.3) can be written as

$$\tilde{x}(k+1) = \left(\tilde{A} + \sum_{i=1}^p v_i(k) \tilde{A}_i \right) \tilde{x}(k) + \tilde{w}(k), \quad (4.1)$$

where $\tilde{x}(k) \triangleq [x^T(k), x_e^T(k)]^T$. To analyze (4.1) it is useful to define the second-moment matrix

$$\tilde{Q}(k) = E[\tilde{x}(k)\tilde{x}^T(k)]. \quad (4.2)$$

It follows from (4.1) and (4.2) that $\tilde{Q}(k)$ must satisfy

$$\tilde{Q}(k+1) = \tilde{A}\tilde{Q}(k)\tilde{A}^T + \sum_{i=1}^P \tilde{A}_i\tilde{Q}(k)\tilde{A}_i^T + \tilde{V}. \quad (4.3)$$

Lemma 4.1. A_e is asymptotically stable if and only if

$$\mathcal{A} \triangleq \tilde{A} \otimes \tilde{A} + \sum_{i=1}^P \tilde{A}_i \otimes \tilde{A}_i$$

is asymptotically stable.

Proof. The result follows from properties of the Kronecker product applied to partitioned matrices. See [36] for details. \square

Hence \mathcal{A} stable assures

$$\tilde{Q} \triangleq \lim_{k \rightarrow \infty} E[\tilde{x}(k)\tilde{x}^T(k)]$$

exists. Furthermore, \tilde{Q} and its nonnegative-definite dual \tilde{P} are the unique solutions of the modified Lyapunov equations

$$\tilde{Q} = \tilde{A}\tilde{Q}\tilde{A}^T + \sum_{i=1}^P \tilde{A}_i\tilde{Q}\tilde{A}_i^T + \tilde{V}, \quad (4.4)$$

$$\tilde{P} = \tilde{A}^T\tilde{P}\tilde{A} + \sum_{i=1}^P \tilde{A}_i^T\tilde{P}\tilde{A}_i + \tilde{R}. \quad (4.5)$$

Partition $(n+n_e) \times (n+n_e)$ \tilde{Q} , \tilde{P} into $n \times n$, $n \times n_e$, and $n_e \times n_e$ subblocks as

$$\tilde{Q} = \begin{bmatrix} Q_1 & Q_{12} \\ Q_{12}^T & Q_2 \end{bmatrix}, \quad \tilde{P} = \begin{bmatrix} P_1 & P_{12} \\ P_{12}^T & P_2 \end{bmatrix},$$

and define the $n \times n$ nonnegative-definite matrices

$$\begin{aligned} Q &\triangleq Q_1 - Q_{12}Q_2^{-1}Q_{12}^T, & \hat{Q} &\triangleq Q_{12}Q_2^{-1}Q_{12}^T, & P &\triangleq P_1 - P_{12}P_2^{-1}P_{12}^T, & \hat{P} &\triangleq P_{12}P_2^{-1}P_{12}^T, \\ \hat{Q} &\triangleq A\hat{Q}A^T + Q_sV_{2s}^{-1}Q_s^T, & \hat{P} &\triangleq (A - Q_sV_{2s}^{-1}C)^T\hat{P}(A - Q_sV_{2s}^{-1}C) + (L - D_eC)^TR(L - D_eC), \end{aligned}$$

where $\tau\hat{Q}\tau^T$ is replaced by \hat{Q} in Q_s and V_{2s} and the $n_e \times n$, $n_e \times n_e$, $n_e \times n$ matrices

$$G \triangleq Q_2^{-1}Q_{12}^T, \quad M \triangleq Q_2P_2, \quad \Gamma \triangleq -P_2^{-1}P_{12}^T.$$

To minimize (3.5) subject to the constraint (4.4), form the Lagrangian

$$\mathcal{L}(A_e, B_e, C_e, D_e, \tilde{Q}, \tilde{P}, \lambda) \triangleq \text{tr} \left[\lambda J(A_e, B_e, C_e, D_e) + \left(\tilde{A}\tilde{Q}\tilde{A}^T + \sum_{i=1}^P \tilde{A}_i\tilde{Q}\tilde{A}_i^T + \tilde{V} - \tilde{Q} \right) \tilde{P} \right],$$

where the Lagrange multipliers $\lambda \geq 0$ and $\tilde{P} \in \mathbf{R}^{(n+n_e) \times (n+n_e)}$ are not both zero. Setting $\partial \mathcal{L} / \partial \tilde{Q} = 0$, $\lambda = 0$

implies $\tilde{P} = 0$ since $(A_e, B_e, C_e, D_e) \in \mathcal{S}^+$. Hence, without loss of generality set $\lambda = 1$. Thus the stationarity conditions are given by

$$\frac{\partial \mathcal{L}}{\partial \tilde{P}} = \tilde{A} \tilde{Q} \tilde{A}^T + \sum_{i=1}^p \tilde{A}_i \tilde{Q} \tilde{A}_i^T + \tilde{V} - \tilde{Q} = 0, \quad (4.6)$$

$$\frac{\partial \mathcal{L}}{\partial \tilde{Q}} = \tilde{A}^T \tilde{P} \tilde{A} + \sum_{i=1}^p \tilde{A}_i^T \tilde{P} \tilde{A}_i + \tilde{R} - \tilde{P} = 0, \quad (4.7)$$

$$\frac{\partial \mathcal{L}}{\partial A_e} = P_{12}^T A Q_{12} + P_2 B_e C Q_{12} + P_2 A_e Q_2 = 0, \quad (4.8)$$

$$\frac{\partial \mathcal{L}}{\partial B_e} = P_{12}^T Q_s + P_2 B_e V_{2s} = 0, \quad (4.9)$$

$$\frac{\partial \mathcal{L}}{\partial C_e} = -R L Q_{12} + R D_e C Q_{12} + R C_e Q_2 = 0, \quad (4.10)$$

$$\frac{\partial \mathcal{L}}{\partial D_e} = D_e V_{2s} - L Q C^T = 0. \quad (4.11)$$

Expanding (4.6) and (4.7) yields

$$A(Q + \hat{Q})A^T + \sum_{i=1}^p A_i(Q + \hat{Q})A_i^T + V_1 - Q - \hat{Q} = 0, \quad (4.12)$$

$$[A\hat{Q}A^T + Q_s V_{2s}^{-1} Q_s^T - \hat{Q}] \Gamma^T = 0, \quad (4.13)$$

$$\Gamma [A\hat{Q}A^T + Q_s V_{2s}^{-1} Q_s^T - \hat{Q}] \Gamma^T = 0, \quad (4.14)$$

$$[(A - Q_s V_{2s}^{-1} C)^T \hat{P} (A - Q_s V_{2s}^{-1} C) + (L - D_e C)^T R (L - D_e C) - \hat{P}] G^T = 0, \quad (4.15)$$

$$G [(A - Q_s V_{2s}^{-1} C)^T \hat{P} (A - Q_s V_{2s}^{-1} C) + (L - D_e C)^T R (L - D_e C) - \hat{P}] G^T = 0. \quad (4.16)$$

Note that the (1, 1) subblock of equation (4.7) which characterizes P_1 has been omitted from the above equations since the estimator gains are independent of P_1 .

Using (4.8)–(4.11) we obtain (3.10)–(3.13). Using (4.12) + $G^T \Gamma (4.13) G - (4.13) G - ((4.13) G)^T$ and $G^T \Gamma (4.13) G - (4.13) G - ((4.13) G)^T + (4.13) - (4.13)$ yields (3.14) and (3.15). Using $\Gamma^T G (4.15) \Gamma - (4.15) \Gamma - ((4.15) \Gamma)^T + (4.15) - (4.15)$ yields (3.16). Finally, $\Gamma (4.13) - (4.14)$ or $G (4.15) - (4.16)$ yields $\Gamma G^T = I_{n_e}$. \square

References

- [1] D.S. Bernstein and D.C. Hyland, The optimal projection equations for reduced-order state estimation, *IEEE Trans. Automat. Control* **30** (1985) 583–585.
- [2] D.S. Bernstein, L.D. Davis, S.W. Greeley and D.C. Hyland, The optimal projection equations for reduced-order, discrete-time modelling, estimation and control, *Proc. 24th IEEE Conf. Decision and Control*, Fort Lauderdale, FL (Dec. 1985) pp. 573–578.
- [3] D.S. Bernstein, L.D. Davis and D.C. Hyland, The optimal projection equations for reduced-order, discrete-time modelling, estimation and control, *J. Guid. Control Dyn.* **9** (1986) 288–293.
- [4] C.S. Sims and J.L. Melsa, A survey of specific optimal techniques in control and estimation, *Internat. J. Control* **12** (1971) 299–308.
- [5] T.E. Fortman and D. Williams, Design of low-order observers for linear feedback control laws, *IEEE Trans. Automat. Control* **17** (1972) 301–308.
- [6] C.S. Sims, An algorithm for estimating a portion of a state vector, *IEEE Trans. Automat. Control* **19** (1974) 391–393.
- [7] R.B. Asher, K.D. Herring and J.C. Ryles, Bias variance and estimation error in reduced-order filters, *Automatica* **12** (1976) 289–600.

- [8] J.I. Galdos and D.E. Gustafson, Information and distortion in reduced-order filter design, *IEEE Trans. Inform. Theory* **23** (1977) 183–194.
- [9] F.W. Fairman, Reduced-order state estimators for discrete-time stochastic systems, *IEEE Trans. Automat. Control* **22** (1977) 673–675.
- [10] F.W. Fairman, On stochastic observer estimators for continuous-time systems, *IEEE Trans. Automat. Control* **22** (1977) 874–876.
- [11] C.S. Sims and R.B. Asher, Optimal and suboptimal results in full and reduced-order linear filtering, *IEEE Trans. Automat. Control* **23** (1978) 469–472.
- [12] C.S. Sims and L.G. Stotts, Linear discrete reduced-order filtering, *Proc. IEEE Conf. Decision and Control* (1979) pp. 1172–1177.
- [13] D.A. Wilson and R.N. Mishra, Design of low order estimators using reduced models, *Internat. J. Control* **23** (1979) 447–456.
- [14] F.W. Fairman and R.D. Gupta, Design of multifunctional reduced-order observers, *Internat. J. Systems Sci.* **11** (1980) 1083–1094.
- [15] U.V. Dombrovskii, Method of synthesizing suboptimal filters of reduced order for digital linear dynamic systems, *Automat. Remote Control* **43** (1982) 1483–1489.
- [16] T. Hinamoto and F.W. Fairman, Reduced order observer design for a linear map of the state, *J. Franklin Inst.* **314** (1982) 95–108.
- [17] C.S. Sims, Reduced-order modelling and filtering, in: C.T. Leondes, Ed., *Control and Dynamic Systems* Vol. 18 (Academic Press, New York, 1982) pp. 55–103.
- [18] P.J. McLane, Optimal linear filtering for linear systems with state-dependent noise, *Internat. J. Control* **10** (1969) 41–51.
- [19] Y. Sunahara and K. Yamashita, An approximate method of estimation for non-linear dynamical systems with state-dependent noise, *Internat. J. Control* **11** (1970) 957–972.
- [20] D.E. Gustafson and J.L. Speyer, Linear minimum variance filters applied to carrier tracking, *IEEE Trans. Automat. Control* **21** (1976) 65–73.
- [21] R.B. Asher and C.S. Sims, Reduced-order filtering with state dependent noise, *Proc. Joint Amer. Control Conf.* (1978).
- [22] C.S. Sims, Discrete reduced-order filtering with state-dependent noise, *Proc. Joint Autom. Control Conf.* (1980).
- [23] H.V. Panossian and C.T. Leondes, Observers for optimal estimation of the state of linear stochastic discrete systems, *Internat. J. Control* **37** (1983) 645–655.
- [24] M.J. Grimble, Wiener and Kalman filters for systems with random parameters, *IEEE Trans. Automat. Control* **29** (1984) 552–554.
- [25] H. Heffes, The effect of erroneous models on the Kalman filter response, *IEEE Trans. Automat. Control* **11** (1966) 541–543.
- [26] J.A. D'Appolito and C.E. Hutchinson, Low sensitivity filter for state estimation in the presence of large parameter uncertainties, *IEEE Trans. Automat. Control* **14** (1969) 310–312.
- [27] D.G. Lainotis and F.L. Sims, Sensitivity analysis of discrete Kalman filters, *Internat. J. Control* **12** (1970) 657–669.
- [28] J.M. Morris, The Kalman filter: A robust estimator for some classes of linear quadratic problems, *IEEE Trans. Inform. Theory* **22** (1976) 526–534.
- [29] C.J. Masreliez and R.D. Martin, Robust Bayesian estimation for the linear model and robustifying the Kalman filter, *IEEE Trans. Automat. Control* **22** (1977) 361–371.
- [30] M. Toda and R.V. Patel, Bounds on estimation error of discrete-time filters under modelling uncertainty, *IEEE Trans. Automat. Control* **25** (1980) 1115–1121.
- [31] R.T. Stefani, Reducing the sensitivity to parameter variations of a minimum-order reduced-order observer, *Internat. J. Control* **35** (1982) 983–995.
- [32] D.S. Bernstein and D.C. Hyland, The optimal projection/maximum entropy approach to designing low-order, robust controllers for flexible structures, *Proc. 24th IEEE Conf. Decision and Control*, Fort Lauderdale, FL (Dec. 1985) pp. 745–752.
- [33] D.S. Bernstein, L.D. Davis, S.W. Greeley and D.C. Hyland, Numerical solution of the optimal projection/maximum entropy design equations for low-order, robust controller design, *Proc. 24th IEEE Conf. Decision and Control*, Fort Lauderdale, FL (Dec. 1985) pp. 1795–1798.
- [34] D.S. Bernstein and S.W. Greeley, Robust controller synthesis using the maximum entropy design equations, *IEEE Trans. Automat. Control* **31** (1986) 362–364.
- [35] D.S. Bernstein and S.W. Greeley, Robust output-feedback stabilization: Deterministic and stochastic perspectives, *Proc. Amer. Control Conf.*, Seattle, WA (June 1986) pp. 1818–1826.
- [36] D.S. Bernstein and D.C. Hyland, The optimal projection equations for reduced-order modelling, estimation and control of linear systems with Stratonovich multiplicative white noise, submitted.
- [37] D.S. Bernstein, Robust static and dynamic output-feedback stabilization: Deterministic and stochastic perspectives, submitted.
- [38] D.S. Bernstein, Robust stability and performance via the extended optimal projection equations for fixed-order dynamic compensation, submitted.
- [39] J.W. Brewer, Kronecker products and matrix calculus in System theory, *IEEE Trans. Circuits and Systems* **25** (1978) 772–781.
- [40] C.R. Rao and S.K. Mitra, *Generalized Inverse of Matrices and Its Applications* (John Wiley and Sons, New York, 1971).

Optimal projection equations for discrete-time fixed-order dynamic compensation of linear systems with multiplicative white noise

DENNIS S. BERNSTEIN† and WASSIM M. HADDAD‡

The optimal projection equations for discrete-time reduced-order dynamic compensation are generalized to include the effects of state-, control- and measurement-dependent noise. In addition, the discrete-time static output feedback problem with multiplicative disturbances is considered. For both problems, the design equations are presented in a concise, unified manner to facilitate their accessibility for developing numerical algorithms for practical applications.

Notation and definitions

| | |
|--------------------------------|---|
| $R, R^{r \times s}, R^r, E$ | real numbers, $r \times s$ real matrices, $R^{r \times l}$, expectation |
| $I_n, (\)^T$ | $n \times n$ identity, transpose |
| \otimes | Kronecker product |
| τ_\perp | $I_n - \tau, \tau \in R^{n \times n}$ |
| asymptotically | |
| stable matrix | matrix with eigenvalues in the open unit disc |
| non-negative-semisimple matrix | semisimple (diagonalizable) matrix with non-negative eigenvalues |
| non-negative-definite matrix | symmetric matrix with non-negative eigenvalues |
| positive-definite matrix | symmetric matrix with positive eigenvalues |
| n, m, l, n_c, p | positive integers, $1 \leq n_c \leq n$ |
| x, x_c | n, n_c -dimensional vectors |
| u, y | m, l -dimensional vectors |
| $A, A_i; B, B_i; C, C_i$ | $n \times n$ matrices, $n \times m$ matrices, $l \times n$ matrices, $i = 1, \dots, p$ |
| A_c, B_c, C_c, D_c | $n_c \times n_c, n_c \times l, m \times n_c, m \times l$ matrices |
| k | discrete-time index $1, 2, \dots$ |
| $v_i(k)$ | unit variance white noise, $i = 1, \dots, p$ |
| $w_1(k), w_2(k)$ | n -dimensional, l -dimensional white noise processes |
| V_1, V_2 | $n \times n$ covariance of $w_1, l \times l$ covariance of w_2 ; $V_1 \geq 0, V_2 \geq 0$ |
| V_{12} | $n \times l$ cross-covariance of w_1, w_2 |
| R_1, R_2 | state and control weightings; $R_1 \geq 0, R_2 \geq 0$ |
| R_{12} | $n \times m$ cross weighting; $R_1 - R_{12}R_2^{-1}R_{12}^T \geq 0$ |
| \tilde{A}, \tilde{A}_i | $A + BD_cC, A_i + B_iD_cC + BD_cC_i, i = 1, \dots, p$ |
| \tilde{w} | $w_1 + BD_cw_2 + \sum_{i=1}^p B_iD_cw_{2i}$ |

Received 23 July 1986.

† Harris Corporation, Government Aerospace Systems Division, Melbourne, FL 32902, U.S.A.

‡ Department of Mechanical Engineering, Florida Institute of Technology, Melbourne, FL 32901, U.S.A.

$$\begin{aligned}
\tilde{V} & V_1 + V_{12}D_c^T B^T + BD_c V_{12}^T + BD_c V_2 D_c^T B^T + \sum_{i=1}^p B_i D_c V_2 D_c^T B_i^T \\
\tilde{R} & R_1 + R_{12}D_c C + C^T D_c^T R_{12}^T + C^T D_c^T R_2 D_c C + \sum_{i=1}^p C_i^T D_c^T R_2 D_c C_i \\
\tilde{A}, \tilde{A}_i & \begin{bmatrix} \tilde{A} & BC_c \\ B_c C & A_c \end{bmatrix}, \quad \begin{bmatrix} \tilde{A}_i & B_i C_c \\ B_c C_i & 0 \end{bmatrix} \\
\tilde{w} & \begin{bmatrix} \tilde{w} \\ B_c w_2 \end{bmatrix} \\
\tilde{V} & \begin{bmatrix} \tilde{V} & V_{12}B_c^T + BD_c V_2 B_c^T \\ B_c V_{12}^T + B_c V_2 D_c^T B^T & B_c V_2 B_c^T \end{bmatrix} \\
\tilde{R} & \begin{bmatrix} \tilde{R} & R_{12}C_c + C^T D_c^T R_2 C_c \\ C_c^T R_{12}^T + C_c^T R_2 D_c C & C_c^T R_2 C_c \end{bmatrix} \\
Z_{(i,j)} & (i,j) \text{ element of matrix } Z \\
\text{tr } Z & \text{trace of square matrix } Z \\
\rho(Z) & \text{rank of matrix } Z \\
E_i & \text{matrix with unity in the } (i,i) \text{ position and zeros elsewhere} \\
\Pi_i(\psi) & \psi E_i \psi^{-1} \text{ (unit-rank eigenprojection)}
\end{aligned}$$

For arbitrary $n \times n$ $Q, P, \hat{Q}, \hat{P}, \tau$ define:

$$\begin{aligned}
V_{2s} &\triangleq V_2 + CQC^T + \sum_{i=1}^p C_i QC_i^T, & R_{2s} &\triangleq R_2 + B^T PB + \sum_{i=1}^p B_i^T P B_i \\
Q_s &\triangleq AQC^T + V_{12} + \sum_{i=1}^p A_i QC_i^T, & P_s &\triangleq B^T PA + R_{12}^T + \sum_{i=1}^p B_i^T P A_i \\
Q_{s1} &\triangleq V_{12} + \sum_{i=1}^p A_i QC_i^T, & P_{s1} &\triangleq R_{12}^T + \sum_{i=1}^p B_i^T P A_i \\
\hat{V}_{2s} &\triangleq V_2 + CQC^T + \sum_{i=1}^p C_i (Q + \tau \hat{Q} \tau^T) C_i^T, & \hat{R}_{2s} &\triangleq R_2 + B^T PB + \sum_{i=1}^p B_i^T (P + \tau^T \hat{P} \tau) B_i \\
\hat{Q}_s &\triangleq AQC^T + V_{12} + \sum_{i=1}^p A_i (Q + \tau \hat{Q} \tau^T) C_i^T, & \hat{P}_s &\triangleq B^T PA + R_{12}^T + \sum_{i=1}^p B_i^T (P + \tau^T \hat{P} \tau) A_i \\
\hat{Q}_{s1} &\triangleq V_{12} + \sum_{i=1}^p A_i (Q + \tau \hat{Q} \tau^T) C_i^T, & \hat{P}_{s1} &\triangleq R_{12}^T + \sum_{i=1}^p B_i^T (P + \tau^T \hat{P} \tau) A_i
\end{aligned}$$

1. Introduction

Hyland and Bernstein (1984) showed that the first-order necessary conditions for quadratically optimal continuous-time fixed-order dynamic compensation can be transformed into a coupled system of four matrix equations (two modified Riccati equations and two modified Lyapunov equations). The coupling is due to the presence of an oblique projection (idempotent matrix) which arises as a rigorous consequence

of optimality. This formulation provides a generalization of classical LQG control theory, since in the full-order case the projection becomes the identity matrix, the modified Lyapunov equations drop out, and the modified Riccati equations reduce to the usual LQG equations. Coupling via the optimal projection implies that sequential reduced-order design procedures consisting of either model reduction followed by controller design or controller design followed by controller reduction are generally suboptimal. Furthermore, the coupled structure of the equations yields the insight that in the reduced-order case there is no longer separation between the operations of state estimation and state-estimate feedback, i.e. the certainty equivalence principle breaks down.

The above developments for the continuous-time problem have, moreover, been carried out by Bernstein, Davis and Hyland (1986) in a discrete-time setting. As in the continuous-time case, the optimal reduced-order compensator is characterized by a pair of modified Riccati equations and a pair of modified Lyapunov equations coupled by an oblique projection. Furthermore, because of the discrete-time setting it is now possible to permit static feedthrough gains in both the full- and reduced-order controller designs. As pointed out by Hyland and Bernstein (1984), non-singular control weighting and measurement noise in the continuous-time case permit only a purely dynamic (strictly proper) controller. Note that this is precisely the case in continuous-time LQG theory, which yields strictly proper feedback controllers.

An immediate application of the discrete-time results is a rigorous treatment of the linear-quadratic sampled-data reduced-order dynamic-compensation problem (Bernstein, Davis and Greeley 1986). By explicitly accounting for real-time computational delay in the feedback loop, the sampled-data control-design problem can be transformed into an equivalent discrete-time problem. The dimension of the equivalent discrete-time system, however, is augmented by the available measurements which are treated as delay states. The optimal projection equations for discrete-time fixed-order dynamic compensation can thus be used to obtain controllers of tractably low dimension in spite of dimension augmentation.

Design considerations concerning stability and performance robustness with respect to unknown parameter variations can also be incorporated into the fixed-order dynamic-compensation design process. This can be accomplished by introducing white noise into the plant via the imperfectly known parameters (Bernstein and Hyland 1985, Bernstein and Greeley 1986 a). Intuitively speaking, the quadratically optimal feedback controller designed in the presence of such disturbances is automatically desensitized to actual parameter variations. As shown by Bernstein and Greeley (1986 b), the modification of the closed-loop covariance equation due to multiplicative noise can be used to guarantee robust stability and performance by means of a Lyapunov function and a performance bound.

An interesting aspect of the design equations for the multiplicative noise model is the breakdown of the separation principle even in the full-order case. That is, even when coupling due to the oblique projection is absent, coupling due to stochastic effects remains. This is a graphic portrayal of observations made previously, e.g. by Gustafson and Speyer (1975). An alternative, apparently suboptimal approach involving certainty equivalent controllers for guaranteeing stochastic stability was developed by Yaz (1986).

The purpose of the present paper is to extend the optimal projection equations for fixed-order discrete-time dynamic compensation given by Bernstein, Davis and Hyland (1986) to include the effects of state-, control- and measurement-dependent

white noise. The main result (Theorem 3.1) presents the necessary conditions for optimality as a system of four matrix equations (two modified discrete-time Riccati equations and two modified discrete-time Lyapunov equations) coupled by both the optimal projection and stochastic effects. For the sake of completeness, the optimality conditions for discrete-time static output feedback are given by Theorem 2.1.

2. Static output feedback

2.1. Discrete-time static output-feedback problem

Given the controlled system

$$x(k+1) = \left(A + \sum_{i=1}^p v_i(k) A_i \right) x(k) + \left(B + \sum_{i=1}^p v_i(k) B_i \right) u(k) + w_1(k) \quad (1)$$

$$y(k) = \left(C + \sum_{i=1}^p v_i(k) C_i \right) x(k) + w_2(k) \quad (2)$$

where $k = 1, 2, \dots$, determine D_c such that the static output feedback law

$$u(k) = D_c y(k) \quad (3)$$

minimizes the performance criterion

$$J \triangleq \lim_{k \rightarrow \infty} \mathbf{E} [x^T(k) R_1 x(k) + 2x^T(k) R_{12} u(k) + u^T(k) R_2 u(k)] \quad (4)$$

Using the notation given at the beginning of this paper, the closed-loop system (1)–(3) can be written as

$$x(k+1) = \left(\tilde{A} + \sum_{i=1}^p v_i(k) \tilde{A}_i \right) x(k) + \tilde{w}(k) \quad (5)$$

Define the second-moment matrix

$$Q(k) = \mathbf{E} [x(k)x^T(k)] \quad (6)$$

satisfying

$$Q(k+1) = \tilde{A} Q(k) \tilde{A}^T + \sum_{i=1}^p \tilde{A}_i Q(k) \tilde{A}_i^T + \tilde{V} \quad (7)$$

To consider the steady state, we restrict our consideration to the set of second-moment stabilizing gains

$$\mathbf{S} \triangleq \left\{ D_c : \tilde{A} \otimes \tilde{A} + \sum_{i=1}^p \tilde{A}_i \otimes \tilde{A}_i \text{ is asymptotically stable} \right\}$$

The requirement $D_c \in \mathbf{S}$ implies the existence of the steady-state closed-loop state covariance $Q \triangleq \lim_{k \rightarrow \infty} Q(k)$. Furthermore, Q and its non-negative-definite dual P are unique solutions of the modified discrete-time Lyapunov equations

$$Q = \tilde{A} Q \tilde{A}^T + \sum_{i=1}^p \tilde{A}_i Q \tilde{A}_i^T + \tilde{V} \quad (8)$$

$$P = \tilde{A}^T P \tilde{A} + \sum_{i=1}^p \tilde{A}_i^T P \tilde{A}_i + \tilde{R} \quad (9)$$

An additional technical requirement is that D_c be confined to the set

$$\mathbf{S}^+ \triangleq \{D_c \in \mathbf{S}: V_{2s} > 0 \text{ and } R_{2s} > 0\}$$

In order to obtain closed-form expressions for extremal values of the closed-loop control gains, the static- and dynamic-compensation problems require the technical assumption

$$[B_i \neq 0 \Rightarrow C_i = 0], \quad i = 1, \dots, p \quad (10)$$

i.e. for each $i \in \{1, \dots, p\}$, B_i and C_i are not both non-zero. Essentially, (10) expresses the condition that the control-dependent and measurement-dependent noises are independent. There are no constraints, however, on correlation with the state-dependent noise. By optimizing (4) with respect to D_c and manipulating (8) and (9), we obtain the following result.

Theorem 2.1

Suppose $D_c \in \mathbf{S}^+$ solves the discrete-time static output-feedback problem. Then there exist $n \times n$ Q , $P \geq 0$ such that

$$D_c = -R_{2s}^{-1} [B^T P A Q C^T + P_{s1} Q C^T + B^T P Q_{s1}] V_{2s}^{-1} \quad (11)$$

and such that Q and P satisfy

$$Q = A Q A^T + V_1 + \sum_{i=1}^p [(A_i + B_i D_c C) Q (A_i + B_i D_c C)^T + B_i D_c V_{2s} D_c^T B_i^T] \\ + (Q_s + B D_c V_{2s}) V_{2s}^{-1} (Q_s + B D_c V_{2s})^T - Q_s V_{2s} Q_s^T \quad (12)$$

$$P = A^T P A + R_1 + \sum_{i=1}^p [(A_i + B D_c C_i)^T P (A_i + B D_c C_i) + C_i^T D_c^T R_{2s} D_c C_i] \\ + (P_s + R_{2s} D_c C)^T R_{2s}^{-1} (P_s + R_{2s} D_c C) - P_s^T R_{2s} P_s \quad (13)$$

3. Dynamic output feedback

We now expand the formulation of the static problem to include a dynamic compensator.

3.1. Discrete-time dynamic output-feedback problem

Given the controlled system (1), (2), determine A_c , B_c , C_c , D_c such that the dynamic output-feedback law

$$x_c(k+1) = A_c x_c(k) + B_c y(k) \quad (14)$$

$$u(k) = C_c x_c(k) + D_c y(k) \quad (15)$$

minimizes the performance criterion (4).

We restrict our attention to the second-moment-stabilizing controllers

$$\mathbf{S} \triangleq \left\{ (A_c, B_c, C_c, D_c): \bar{A} \otimes \bar{A} + \sum_{i=1}^p \bar{A}_i \otimes \bar{A}_i \text{ is asymptotically stable and} \right. \\ \left. (A_c, B_c, C_c) \text{ is minimal} \right\}$$

which implies the existence of $\bar{Q} \triangleq \lim_{k \rightarrow \infty} E[\bar{x}(k) \bar{x}^T(k)]$, where $\bar{x}(k) \triangleq [x^T(k), x_c^T(k)]^T$.

Furthermore, \tilde{Q} and its non-negative-definite dual \tilde{P} are the unique solutions to the modified discrete-time Lyapunov equations

$$\tilde{Q} = \tilde{A}\tilde{Q}\tilde{A}^T + \sum_{i=1}^p \tilde{A}_i\tilde{Q}\tilde{A}_i^T + \tilde{V} \quad (16)$$

$$\tilde{P} = \tilde{A}^T\tilde{P}\tilde{A} + \sum_{i=1}^p \tilde{A}_i^T\tilde{P}\tilde{A}_i + \tilde{R} \quad (17)$$

An additional technical assumption is that (A_c, B_c, C_c, D_c) be confined to the set

$$\mathbf{S}^+ \triangleq \{(A_c, B_c, C_c, D_c) \in \mathbf{S} : \hat{R}_{2s} > 0 \text{ and } \hat{V}_{2s} > 0\}$$

The following lemma is required.

Lemma 3.1

Let $\tau \in \mathbf{R}^{n \times n}$. Then

$$\tau^2 = \tau \quad (18)$$

$$\rho(\tau) = n_c \quad (19)$$

if and only if there exist $G, \Gamma \in \mathbf{R}^{n_c \times n}$ such that

$$G^T \Gamma = \tau \quad (20)$$

$$\Gamma G^T = I_{n_c} \quad (21)$$

Proof

See Bernstein, Davis and Hyland (1986).

For convenience call G and Γ satisfying (20) and (21) a *projective factorization* of τ . Furthermore, for $n \times n$ non-negative-definite matrices \hat{Q}, \hat{P} , define the set of *contragrediently diagonalizing transformations* (see Rao and Mitra 1971, p. 123)

$$\mathbf{D}(\hat{Q}, \hat{P}) \triangleq \{\psi \in \mathbf{R}^{n \times n} : \psi^{-1} \hat{Q} \psi^{-T} \text{ and } \psi^T \hat{P} \psi \text{ are diagonal}\}$$

Theorem 3.1

Suppose $(A_c, B_c, C_c, D_c) \in \mathbf{S}^+$ solves the discrete-time dynamic output-feedback problem. Then there exist $n \times n$ $Q, P, \hat{Q}, \hat{P} \geq 0$ such that

$$A_c = \Gamma[A - B\hat{R}_{2s}^{-1}\hat{P}_s - \hat{Q}_s\hat{V}_{2s}^{-1}C - BD_cC]G^T \quad (22)$$

$$B_c = \Gamma[\hat{Q}_s\hat{V}_{2s}^{-1} + BD_c] \quad (23)$$

$$C_c = -[\hat{R}_{2s}^{-1}\hat{P}_s + D_cC]G^T \quad (24)$$

$$D_c = -\hat{R}_{2s}^{-1}[B^T P A Q C^T + \hat{P}_{s1} Q C^T + B^T P \hat{Q}_{s1}] \hat{V}_{2s}^{-1} \quad (25)$$

and such that Q, P, \hat{Q}, \hat{P} satisfy

$$\begin{aligned} Q = & A Q A^T + V_1 + \tau_1 \hat{Q} \tau_1^T + \sum_{i=1}^p [(A_i - B_i \hat{R}_{2s}^{-1} \hat{P}_s) \tau \hat{Q} \tau^T (A_i - B_i \hat{R}_{2s}^{-1} \hat{P}_s)^T \\ & + (A_i + B_i D_c C) Q (A_i + B_i D_c C)^T + B_i D_c V_2 D_c^T B_i^T] - \hat{Q}_s \hat{V}_{2s}^{-1} \hat{Q}_s^T \end{aligned} \quad (26)$$

$$P = A^T P A + R_1 + \tau_1^T \hat{P} \tau_1 + \sum_{i=1}^p [(A_i - \hat{Q}_s \hat{V}_{2s}^{-1} C_i)^T \tau^T \hat{P} \tau (A_i - \hat{Q}_s \hat{V}_{2s}^{-1} C_i) + (A_i + B D_c C_i)^T P (A_i + B D_c C_i) + C_i^T D_c^T R_2 D_c C_i] - \hat{P}_s^T \hat{R}_{2s}^{-1} \hat{P}_s \quad (27)$$

$$\hat{Q} = (A - B \hat{R}_{2s}^{-1} \hat{P}_s) \tau \hat{Q} \tau^T (A - B \hat{R}_{2s}^{-1} \hat{P}_s)^T + (\hat{Q}_s + B D_c \hat{V}_{2s}) \hat{V}_{2s}^{-1} (\hat{Q}_s + B D_c \hat{V}_{2s})^T \quad (28)$$

$$\hat{P} = (A - \hat{Q}_s \hat{V}_{2s}^{-1} C)^T \tau^T \hat{P} \tau (A - \hat{Q}_s \hat{V}_{2s}^{-1} C) + (\hat{P}_s + \hat{R}_{2s} D_c C)^T \hat{R}_{2s}^{-1} (\hat{P}_s + \hat{R}_{2s} D_c C) \quad (29)$$

where

$$\tau \triangleq \sum_{i=1}^{n_c} \Pi_i(\psi) = G^T \Gamma \quad (30)$$

for some $\psi \in \mathbf{D}(\hat{Q}, \hat{P})$ such that $(\psi^{-1} \hat{Q} \hat{P} \psi)_{(i,i)} \neq 0$, $i = 1, \dots, n_c$, and some projective factorization G, Γ of τ .

Remark 3.1

To specialize the result to the strictly proper (no feedthrough) case, merely ignore (25) and set $D_c = 0$ wherever it appears.

Remark 3.2

As previously pointed out by Bernstein, Davis and Hyland (1986), the indeterminacy in specifying the projective factorization G, Γ satisfying (20) and (21) corresponds to an arbitrary choice of internal state-space basis for the design system (A_c, B_c, C_c) .

Remark 3.3

In the full-order case $n_c = n$, the projection τ becomes the identity and (28) and (29) play no role. In this case $G^T \Gamma = \Gamma G^T = I_n$ and thus G and Γ can be chosen to be the identity. Deleting all multiplicative white noise terms corresponding to state-, control- and measurement-dependent disturbances, i.e. $A_i, B_i, C_i = 0$, $i = 1, \dots, p$, and specializing further to the purely dynamic case ($D_c = 0$) yields the standard LQG result. Alternatively, setting $n_c < n$ and deleting the multiplicative noise terms yields the results of Bernstein, Davis and Hyland (1986).

4. Proof of Theorem 3.1

Partition $(n + n_c) \times (n + n_c)$ \tilde{Q}, \tilde{P} into $n \times n$, $n \times n_c$, and $n_c \times n_c$ sub-blocks as

$$\tilde{Q} = \begin{bmatrix} Q_1 & Q_{12} \\ Q_{12}^T & Q_2 \end{bmatrix}, \quad \tilde{P} = \begin{bmatrix} P_1 & P_{12} \\ P_{12}^T & P_2 \end{bmatrix}$$

and define the $n \times n$ non-negative-definite matrices

$$Q \triangleq Q_1 - Q_{12} Q_2^{-1} Q_{12}^T, \quad P \triangleq P_1 - P_{12} P_2^{-1} P_{12}^T$$

$$\hat{Q} \triangleq Q_{12} Q_2^{-1} Q_{12}^T, \quad \hat{P} \triangleq P_{12} P_2^{-1} P_{12}^T$$

$$\hat{Q} \triangleq (A - B \hat{R}_{2s}^{-1} \hat{P}_s) \hat{Q} (A - B \hat{R}_{2s}^{-1} \hat{P}_s)^T + (\hat{Q}_s + B D_c \hat{V}_{2s}) \hat{V}_{2s}^{-1} (\hat{Q}_s + B D_c \hat{V}_{2s})^T$$

$$\hat{P} \triangleq (A - \hat{Q}_s \hat{V}_{2s}^{-1} C)^T \hat{P} (A - \hat{Q}_s \hat{V}_{2s}^{-1} C) + (\hat{P}_s + \hat{R}_{2s} D_c C)^T \hat{R}_{2s}^{-1} (\hat{P}_s + \hat{R}_{2s} D_c C)$$

where $\tau \hat{Q} \tau^T$ and $\tau^T \hat{P} \tau$ in $\hat{Q}, \hat{P}, \hat{V}_{2s}$, and \hat{R}_{2s} are replaced by \hat{Q} and \hat{P} , and the $n_c \times n$,

$n_c \times n_c$, and $n_c \times n$ matrices

$$G \triangleq Q_2^{-1} Q_{12}^T, \quad M \triangleq Q_2 P_2, \quad \Gamma \triangleq -P_2^{-1} P_{12}^T$$

Define the lagrangian

$$L(A_c, B_c, C_c, D_c, \tilde{Q}, \tilde{P}, \lambda) \triangleq \text{tr} \left[\lambda J(A_c, B_c, C_c, D_c) + \left(\tilde{A} \tilde{Q} \tilde{A}^T + \sum_{i=1}^p \tilde{A}_i \tilde{Q} \tilde{A}_i^T + \tilde{V} - \tilde{Q} \right) \tilde{P} \right]$$

where the Lagrange multipliers $\lambda \geq 0$ and $\tilde{P} \in \mathbf{R}^{(n+n_c) \times (n+n_c)}$ are not both zero. Setting $\partial L / \partial \tilde{Q} = 0$, $\lambda = 0$ implies $\tilde{P} = 0$ since $(A_c, B_c, C_c, D_c) \in \mathbf{S}^+$. Hence, without loss of generality, set $\lambda = 1$. Thus the stationarity conditions are given by

$$\frac{\partial L}{\partial \tilde{P}} = \tilde{A} \tilde{Q} \tilde{A}^T + \sum_{i=1}^p \tilde{A}_i \tilde{Q} \tilde{A}_i^T + \tilde{V} - \tilde{Q} = 0 \quad (31)$$

$$\frac{\partial L}{\partial \tilde{Q}} = \tilde{A}^T \tilde{P} \tilde{A} + \sum_{i=1}^p \tilde{A}_i^T \tilde{P} \tilde{A}_i + \tilde{R} - \tilde{P} = 0 \quad (32)$$

$$\frac{\partial L}{\partial A_c} = P_{12}^T A Q_{12} + P_{12}^T B D_c C Q_{12} + P_{12}^T B C_c Q_2 + P_2 A_c Q_2 + P_2 B_c C Q_{12} = 0 \quad (33)$$

$$\frac{\partial L}{\partial B_c} = P_2 B_c \hat{V}_{2s} + P_{12}^T \hat{Q}_s + P_{12}^T B D_c \hat{V}_{2s} = 0 \quad (34)$$

$$\frac{\partial L}{\partial C_c} = \hat{P}_s Q_{12} + \hat{R}_{2s} C_c Q_2 + \hat{R}_{2s} D_c C Q_{12} = 0 \quad (35)$$

$$\frac{\partial L}{\partial D_c} = \hat{R}_{2s} D_c \hat{V}_{2s} + B^T P A Q C^T + \hat{P}_{s1} Q C^T + B^T P \hat{Q}_{s1} = 0 \quad (36)$$

Expanding (31) and (32) yields

$$\begin{aligned} & A Q A^T + (\hat{Q}_s + B D_c \hat{V}_{2s}) \hat{V}_{2s}^{-1} (\hat{Q}_s + B D_c \hat{V}_{2s})^T - \hat{Q}_s \hat{V}_{2s}^{-1} \hat{Q}_s^T \\ & + (A - B \hat{R}_{2s}^{-1} \hat{P}_s) \hat{\hat{Q}} (A - B \hat{R}_{2s}^{-1} \hat{P}_s)^T \\ & + V_1 - Q - \hat{Q} + \sum_{i=1}^p [(A_i - B_i \hat{R}_{2s}^{-1} \hat{P}_s) \hat{\hat{Q}} (A_i - B_i \hat{R}_{2s}^{-1} \hat{P}_s)^T \\ & + (A_i + B_i D_c C) Q (A_i + B_i D_c C)^T + B_i D_c V_2 D_c^T B_i^T] = 0 \end{aligned} \quad (37)$$

$$[(\hat{Q}_s + B D_c \hat{V}_{2s}) \hat{V}_{2s}^{-1} (\hat{Q}_s + B D_c \hat{V}_{2s})^T + (A - B \hat{R}_{2s}^{-1} \hat{P}_s) \hat{\hat{Q}} (A - B \hat{R}_{2s}^{-1} \hat{P}_s)^T - \hat{Q}] \Gamma^T = 0 \quad (38)$$

$$\begin{aligned} & \Gamma [(\hat{Q}_s + B D_c \hat{V}_{2s}) \hat{V}_{2s}^{-1} (\hat{Q}_s + B D_c \hat{V}_{2s})^T \\ & + (A - B \hat{R}_{2s}^{-1} \hat{P}_s) \hat{\hat{Q}} (A - B \hat{R}_{2s}^{-1} \hat{P}_s)^T - \hat{Q}] \Gamma^T = 0 \end{aligned} \quad (39)$$

$$\begin{aligned} & A^T P A + (\hat{P}_s + \hat{R}_{2s} D_c C)^T \hat{R}_{2s}^{-1} (\hat{P}_s + \hat{R}_{2s} D_c C) - \hat{P}_s^T \hat{R}_{2s}^{-1} \hat{P}_s \\ & + (A - \hat{Q}_s \hat{V}_{2s}^{-1} C)^T \hat{\hat{P}} (A - \hat{Q}_s \hat{V}_{2s}^{-1} C) \\ & + R_1 - P - \hat{\hat{P}} + \sum_{i=1}^p [(A_i - \hat{Q}_s \hat{V}_{2s}^{-1} C_i)^T \hat{\hat{P}} (A_i - \hat{Q}_s \hat{V}_{2s}^{-1} C_i) \\ & + (A_i + B D_c C_i)^T P (A_i + B D_c C_i) + C_i^T D_c^T R_2 D_c C_i] = 0 \end{aligned} \quad (40)$$

$$[(\hat{P}_s + \hat{R}_{2s} D_c C)^T \hat{R}_{2s}^{-1} (\hat{P}_s + \hat{R}_{2s} D_c C) + (A - \hat{Q}_s \hat{V}_{2s}^{-1} C)^T \hat{\hat{P}} (A - \hat{Q}_s \hat{V}_{2s}^{-1} C) - \hat{\hat{P}}] G^T = 0 \quad (41)$$

$$G[(\hat{P}_s + \hat{R}_{2s}D_cC)^T \hat{R}_{2s}^{-1}(\hat{P}_s + \hat{R}_{2s}D_cC) + (A - \hat{Q}_s\hat{V}_{2s}^{-1}C)^T \hat{P}(A - \hat{Q}_s\hat{V}_{2s}^{-1}C) - \hat{P}]G^T = 0 \quad (42)$$

Using (33)–(36) we obtain (22)–(25). Using (37) + $G^T\Gamma(38)G - (38)G - (38G)^T$ and $G^T\Gamma(38)G - (38)G - (38G)^T$ yields (26) and (28). Similarly, using (40) + $\Gamma^TG(41)\Gamma - (41)\Gamma - (41\Gamma)^T$ and $\Gamma^TG(41)\Gamma - (41)\Gamma - (41\Gamma)^T$ we obtain (27) and (29). Also, $\Gamma(38) - (39)$ or $G(41) - (42)$ yields $\Gamma G^T = I_{n_c}$ so that $\tau = G^T\Gamma = \tau^2$. Finally, (39) and (42) imply $\hat{Q} = \tau\hat{Q}\tau^T$ and $\hat{P} = \tau^T\hat{P}\tau$. \square

Remark 4.1

An interesting difference between the above discrete-time derivation and the continuous-time derivation of Hyland and Bernstein (1984) is that the explicit gain expressions and the definition of the optimal projection arise in the reverse order.

5. Directions for further research

The principal application of Theorems 2.1 and 3.1 is the sampled-data problem with parameter uncertainties. Although generalization of the results of Bernstein *et al.* (1986) is possible, there appear to be a number of mathematical issues which arise. A related development appears in Tiedemann and De Koning (1984). A more extensive treatment of the results of the present paper can be found in Haddad (1987).

ACKNOWLEDGMENT

Dr. Bernstein was supported in part by the Air Force Office of Scientific Research under contract AFOSR F49620-86-C-0002.

REFERENCES

- BERNSTEIN, D. S., DAVIS, L. D., and GREELEY, S. W., 1986, *I.E.E.E. Trans. autom. Control*, **31**, 859.
- BERNSTEIN, D. S., DAVIS, L. D., and HYLAND, D. C., 1986, *J. Guid. Control Dynam.*, **9**, 288.
- BERNSTEIN, D. S., and GREELEY, S. W., 1986 a, *I.E.E.E. Trans. autom. Control*, **31**, 362; 1986 b, *Proc. American Control Conf.*, Seattle, WA, pp. 1818–1826.
- BERNSTEIN, D. S., and HYLAND, D. C., 1985, *Proc. 24th I.E.E.E. Conf. on Decision and Control*, Fort Lauderdale, FL, pp. 745–752.
- GUSTAFSON, D. E., and SPEYER, J. L., 1975, *J. Spacecraft Rockets*, **12**, 351.
- HADDAD, W. M., 1987, Robust optimal projection control-system synthesis. Ph.D. dissertation, Department of Mechanical Engineering, Florida Institute of Technology.
- HYLAND, D. C., and BERNSTEIN, D. S., 1984, *I.E.E.E. Trans. autom. Control*, **29**, 1034.
- RAO, C. R., and MITRA, S. K., 1971, *Generalized Inverse of Matrices and Its Applications* (New York: Wiley).
- TIEDEMANN, A. R., and DE KONING, W. L., 1984, *Int. J. Control*, **70**, 449.
- YAZ, E., 1986, *I.E.E.E. Trans. autom. Control*, **31**, 586.

INEQUALITIES FOR THE TRACE OF MATRIX EXPONENTIALS*

DENNIS S. BERNSTEIN†

Abstract. Several inequalities involving the trace of matrix exponentials are derived. The Golden-Thompson inequality $\text{tr } e^{A+B} \leq \text{tr } e^A e^B$ for symmetric A and B is obtained as a special case along with the new inequality $\text{tr } e^A e^{A^T} \leq \text{tr } e^{A+A^T}$ for nonnormal A .

Key words. matrix exponential, inequality, trace

AMS(MOS) subject classification. 15

1. Introduction. For $n \times n$ real symmetric matrices A and B , the Golden-Thompson inequality [1]–[5] states that

$$(1.1) \quad \text{tr } e^{A+B} \leq \text{tr } e^A e^B.$$

Reference [5] generalizes (1.1) to allow arbitrary spectral functions in place of the trace operator and provides an overview of applications in which these inequalities arise.

In contrast to (1.1), problems in linear-quadratic optimal feedback control [6] typically involve a performance functional J of the form

$$(1.2) \quad J = \text{tr} \int_0^\infty e^{At} V e^{A^T t} R dt,$$

where V and R denote noise intensity and performance weighting matrices, respectively, and A denotes the linear system dynamics matrix. The form of (1.2) thus suggests inequalities of the form (1.1) involving A and A^T , where A is nonnormal, in place of symmetric A and B . Such inequalities are motivated by robust sampled-data control-design problems which require performance bounds for uncertain system models.

The main result of the present note is the inequality

$$(1.3) \quad \text{tr } e^A e^{A^T} \leq \text{tr } e^{A+A^T}.$$

Rather surprisingly, the sign of the inequality (1.3) is opposite to the sign of (1.1). To understand why this is the case, we derive a series of inequalities which, upon appropriate specialization, yield both (1.1) and (1.3).

2. Inequalities. The following lemma is required. (Let C^T denote the transpose of a matrix C .)

LEMMA 2.1. If $C \in R^{n \times n}$ and r is a positive integer, then

$$(2.1) \quad \text{tr } C^{2r} \leq \text{tr } C^r C^{rT} \leq \text{tr } (C C^T)^r.$$

Proof. The first inequality follows from $\text{tr } (C^r - C^{rT})(C^{rT} - C^r) \geq 0$, while the second follows from a result of K. Fan (see [4, pp. 234, 516]). \square

THEOREM 2.1. If $A, B \in R^{n \times n}$, then

$$(2.2) \quad \begin{aligned} \text{tr } e^{A+B} &\leq \text{tr } e^{(A+B)/2} e^{(A+B)^T/2} \leq \text{tr } e^{(A+A^T)/2} e^{(B+B^T)/2} \\ &\leq \text{tr } e^{(A+A^T)/2} e^{(B+B^T)/2} \leq \frac{1}{2} \text{tr } (e^{A+A^T} + e^{B+B^T}). \end{aligned}$$

$$(2.3) \quad \left. \begin{aligned} \text{tr } e^A e^B \\ \frac{1}{2} \text{tr } (e^{2A} + e^{2B}) \end{aligned} \right\} \leq \frac{1}{2} \text{tr } (e^A e^{A^T} + e^B e^{B^T}) \leq \frac{1}{2} \text{tr } (e^{A+A^T} + e^{B+B^T}).$$

* Received by the editors January 21, 1987; accepted for publication May 11, 1987. This research was supported in part by the Air Force Office of Scientific Research under contract F49620-86-C-0002.

† Harris Corporation, Melbourne, Florida 32902

Proof. Defining $C = e^{A/2r} e^{B/2r}$, (2.1) becomes

$$\operatorname{tr} (e^{A/2r} e^{B/2r})^{2r} \leq \operatorname{tr} (e^{A/2r} e^{B/2r})^r (e^{B^T/2r} e^{A^T/2r})^r \leq \operatorname{tr} (e^{A/2r} e^{B/2r} e^{B^T/2r} e^{A^T/2r})^r.$$

Letting $r \rightarrow \infty$, the exponential product formula [5, p. 60] and its generalization [7, p. 97] yield the first two inequalities of (2.2). The third inequality of (2.2) follows from Corollary 3 of [5] while the fourth inequality of (2.2) follows from

$$0 \leq \operatorname{tr} [e^{(A+A^T)/2} - e^{(B+B^T)/2}]^2.$$

To prove (2.3) note that the upper leftmost inequality follows from $0 \leq \operatorname{tr} (e^A - e^B)(e^A - e^B)^T$. The remaining inequalities in (2.3) follow from $\operatorname{tr} e^{2A} \leq \operatorname{tr} e^A e^{A^T} \leq \operatorname{tr} e^{A+A^T}$, which is a consequence of (2.2) with $B = A$. \square

COROLLARY 2.1. If $A \in R^{n \times n}$, then

$$(2.4) \quad \operatorname{tr} e^{2A} \leq \operatorname{tr} e^A e^{A^T} \leq \operatorname{tr} e^{A+A^T} \leq \frac{n}{2} + \frac{1}{2} \operatorname{tr} e^{2(A+A^T)},$$

$$(2.5) \quad \operatorname{tr} e^{2A} \leq \frac{n}{2} + \frac{1}{2} \operatorname{tr} e^{2A} e^{2A^T} \leq \frac{n}{2} + \frac{1}{2} \operatorname{tr} e^{2(A+A^T)}.$$

If $A, B \in R^{n \times n}$ are symmetric, then

$$(2.6) \quad \operatorname{tr} e^{A+B} \leq \operatorname{tr} e^A e^B \leq \frac{1}{2} \operatorname{tr} (e^{2A} + e^{2B}).$$

Proof. The first two inequalities of (2.4) follow from the first two inequalities of (2.2) with $B = A$. The last inequality of (2.4) follows from the last inequality of (2.2) with $B = 0$ and A replaced by $2A$. Inequalities (2.5) follow from (2.3) with $B = 0$ and A replaced by $2A$ while ignoring the lower leftmost term in (2.3). Finally, (2.6) follows from (2.2). \square

Remark. The second inequality in (2.4) and the first inequality in (2.6) correspond to (1.3) and (1.1), respectively.

3. Additional inequalities. The question immediately arises as to whether any additional inequalities involving the expressions appearing in (2.4) and (2.5) are true. Note that $\operatorname{tr} e^A e^B$ in (2.3) cannot be merged with (2.2) because of the sign reversal between (1.1) and (1.3). It can readily be seen that the only remaining possibilities are

$$(3.1) \quad \operatorname{tr} e^{(A+A^T)/2} e^{(B+B^T)/2} \leq \frac{1}{2} \operatorname{tr} (e^A e^{A^T} + e^B e^{B^T}),$$

$$(3.2) \quad \operatorname{tr} e^{(A+A^T)/2} e^{(B+B^T)/2} \leq \frac{1}{2} \operatorname{tr} (e^{2A} + e^{2B}),$$

$$(3.3) \quad \operatorname{tr} e^A e^B \leq \frac{1}{2} \operatorname{tr} (e^{2A} + e^{2B}).$$

By randomly generating A and B , (3.1) was shown to be false. Since (3.2) implies (3.1), (3.2) must also be false. Furthermore, in the case $B^T = -B$, inequality (3.1), which becomes

$$(3.4) \quad \operatorname{tr} e^{(A+A^T)/2} \leq \frac{n}{2} + \frac{1}{2} \operatorname{tr} e^A e^{A^T},$$

was also shown to be false. Hence (2.4) and (2.5) cannot be merged. Finally, the remaining inequality (3.3) was also shown to be false even when $B = 0$.

Remark. The results of this paper can be generalized to the case in which A and B are complex matrices. Generalization to arbitrary spectral functions [5] remains an area for further research.

Acknowledgment. I wish to thank Scott W. Greeley for carrying out numerical calculations which suggested the results of this paper.

REFERENCES

- [1] S. GOLDEN, *Lower bounds for the Helmholtz function*, Phys. Rev., 137 (1965), pp. B1127-B1128.
- [2] C. J. THOMPSON, *Inequality with applications in statistical mechanics*, J. Math. Phys., 6 (1965), pp. 1812-1813.
- [3] A. LENARD, *Generalization of the Golden-Thompson Inequality* $\text{Tr}(e^A e^B) \geq \text{Tr} e^{A+B}$, Indiana Univ. Math J., 21 (1971), pp. 457-467.
- [4] A. W. MARSHALL AND I. OLKIN, *Inequalities: Theory of Majorization and Its Applications*, Academic Press, New York, 1979.
- [5] J. E. COHEN, S. FRIEDLAND, T. KATO, AND F. P. KELLY, *Eigenvalue inequalities for products of matrix exponentials*, Linear Algebra Appl., 45 (1982), pp. 55-95.
- [6] K. KWAKERNAAK AND R. SIVAN, *Linear Optimal Control Systems*, John Wiley, New York, 1972.
- [7] V. S. VARADARAJAN, *Lie Groups, Lie Algebras, and Their Representations*, Springer-Verlag, Berlin, New York, Heidelberg, 1984.

PROBLEMS AND SOLUTIONS

EDITED BY MURRAY S. KLAMKIN

COLLABORATING EDITORS: CECIL C. ROUSSEAU OTTO G. RUEHR

All problems and solutions should be sent, typewritten in duplicate, to Murray S. Klamkin, Department of Mathematics, University of Alberta, Edmonton, Alberta, Canada T6G 2G1. An asterisk placed beside a problem number indicates that the problem was submitted without solution. Proposers and solvers whose solutions are published will receive 5 reprints of the corresponding problem section. Other solvers will receive just one reprint provided a self-addressed stamped (U.S.A. or Canada) envelope is enclosed. Proposers and solvers desiring acknowledgment of their contributions should include a self-addressed stamped postcard (no stamps necessary outside the U.S.A. and Canada). Solutions should be received by June 30, 1988.

PROBLEMS

Commuting Matrix Exponentials

Problem 88-1, by DENNIS S. BERNSTEIN (Harris Corporation, Melbourne, Florida).*

In feedback control theory for sampled-data systems, the equivalent discrete-time dynamics matrix is given by e^{Ah} , where h is the sample interval and A is the dynamics matrix for the original continuous-time system. When A is perturbed by A_0 (due possibly to some modeling uncertainty), then it is necessary to consider $e^{(A+A_0)h}$. For robust control system design it thus may be of interest to know when $e^{(A+A_0)h}$ can be decomposed into a nominal part involving A and a perturbed part involving A_0 . Analogous questions arise in the study of bilinear control systems of the form

$$\dot{x} = Ax + uBx,$$

where u is a scalar control. In this case the Lie group generated by A and B plays a central role. Again, it is of interest to know how e^{A+B} is related to e^A and e^B , the principal result being the Baker-Campbell-Hausdorff formula. Of course, it is well known that when

$$(1) \quad AB = BA,$$

where A, B are real $n \times n$ matrices, then both

$$(2) \quad e^A e^B = e^B e^A$$

and

$$(3) \quad e^{-A} e^B = e^{B-A}$$

hold. It is less well known that the converse is not true. Specifically, examples are given in [1] which show that (2) may hold while (1) and (3) are violated, and that (3) may hold while (1) is violated. Interestingly, it is stated without proof in [1] that (3) implies (2). Prove this claim or find a counterexample. A copy of [1] is available from the proposer.

REFERENCE

- [1] M. FRECHET, *Les solutions non commutables de l'équation matricielle $e^x e^y = e^{x+y}$* , Rend. Circ. Mat. Palermo 2, (1952), pp. 11-27.

ROBUST STABILITY FOR SAMPLED-DATA CONTROL SYSTEMS

by

Dennis S. Bernstein
Harris Corporation
Government Aerospace Systems Division
MS 22/4848
Melbourne, FL 32902

C.V. Hollot
Department of Electrical
and Computer Engineering
University of Massachusetts
Amherst, MA 01003

ABSTRACT

In this paper we consider the robust stability of a continuous-time system under computer control. The uncertainty is modeled as additive perturbations to the matrices in a continuous-time state space description of the plant. Our methods exploit the resulting exponential-like uncertainty structure in the sampled-data control system and we develop sufficient conditions for such a system to be robustly stable.

1. Introduction

A sampled-data control system consists of a continuous-time plant under computer control; see Figure 1. Generally speaking, if the matrices in a state space description of a continuous-time plant are uncertain, then the resulting closed-loop, discrete-time system possesses an exponential-like uncertainty structure. This is true even if the continuous-time plant has linear uncertainty. Existing methods; e.g., see [1]-[4], are inadequate in analyzing such uncertain discrete-time systems since they do not directly handle these exponential-like structures. Notable exceptions include the conic sector approach in [5], and the stochastic parameter formulation in [6]. Indeed, the present paper was motivated by the approach of [6] which, as shown in [12] can be reinterpreted to yield conditions for deterministic robust stability. Our objective is thus to develop a robust stability test which exploits the specific nonlinear uncertainty structures occurring in sampled-data control systems.

In the sequel, the following notation will be used. For $X \in R^{n \times n}$, X' denotes the transpose of X , while $X > 0$ ($X \geq 0$) means that X is positive definite (positive semi-definite). The spectral radius of X is given by $\rho(X)$. Let \otimes , \oplus and "vec" denote the Kronecker product, Kronecker sum and column stacking operators respectively; see [9]. In addition we shall use "vec⁻¹" to denote the operation of forming a (usually square) matrix from a column vector.

2. Problem Formulation

In this section we state the robust analysis problem for sampled-data control systems using static feedback. To begin, consider the n -dimensional continuous-time plant

$$\begin{aligned}\dot{x}(t) &= (A + \Delta A)x(t) + (B + \Delta B)u(t) \\ y(t) &= Cx(t)\end{aligned}$$

(Σ)

where $A \in R^{n \times n}$, $B \in R^{n \times m}$ and $C \in R^{1 \times n}$ denote nominal state space data and where ΔA and ΔB represent perturbations in A and B respectively. The pair

of suitably dimensioned matrices $(\Delta A, \Delta B)$ belongs to an uncertainty set U given by

$$U = \{(\Delta A, \Delta B): \Delta A = \sum_{i=1}^p \sigma_i A_i \text{ and } \Delta B = \sum_{i=1}^p \sigma_i B_i, \sum_{i=1}^p \sigma_i^2 \leq 1\} \quad (2.1)$$

where A_i and B_i reflect the "structure" of the uncertainty and where σ_i is an uncertain real parameter; note that an uncertain parameter σ_i may appear in both ΔA and ΔB , and it's possible to have $A_i = 0$ and $B_i \neq 0$ or vice-versa.

Now, consider the sampled-data system in Figure 1 with a sampling period of h seconds. We assume perfect synchronization between the A/D (sampler) and the (D/A) (zero-order hold) and ignore finite word-length effects and computational delays. We also assume that a static control law

$$u(kh) = Ky(kh); \quad k = 0, 1, 2, \dots, \quad (2.2)$$

is implemented for some given gain $K \in \mathbb{R}^{m \times l}$.

Our purpose is to analyze the robust stability of this closed-loop system, and to this end we consider its evolution at the sample instances kh by forming the associated discrete-time system

$$x(k+1) = [e^{(A + \Delta A)h} + \int_0^h e^{(A + \Delta A)\tau} d\tau (B + \Delta B)KC]x(k). \quad (2.3)$$

In the above $x(k)$ denotes $x(kh)$; we have abused notation for the sake of conciseness. Given arbitrary $(\Delta A, \Delta B) \in U$, system (2.3) is discrete-time stable if all the eigenvalues of

$$e^{(A + \Delta A)h} + \int_0^h e^{(A + \Delta A)\tau} d\tau (B + \Delta B)KC$$

lie within the open unit disk. Additionally, (2.3) is said to be robustly discrete-time stable if it is discrete-time stable for all $(\Delta A, \Delta B) \in U$.

3. Main Result

We now develop a sufficient condition for the robust stability of (2.3). This condition exploits the exponential structure of the uncertainty; i.e., rather than "overbounding" the uncertainty with an additive model of the form

$$x(k+1) = (\Phi + \Delta\Phi)x(k),$$

our methods treat the exponential structure of the nonlinearity in (2.3) more directly.

To show robust stability we will construct a parameter-independent (independent of the uncertain matrices $(\Delta A, \Delta B)$) quadratic Lyapunov function $V(x)$ for (2.3). Thus, let P be some free, positive-definite symmetric matrix and consider the quadratic Lyapunov candidate

$$V(x) = x'Px, \quad x \in \mathbb{R}^n.$$

System (2.3) is robustly discrete-time stable if

$$\Delta V(\Delta A, \Delta B, P) \stackrel{\Delta}{=} [e^{(A+\Delta A)h} + \int_0^h e^{(A+\Delta A)\tau} d\tau (B+\Delta B)KC]' P [e^{(A+\Delta A)h} + \int_0^h e^{(A+\Delta A)\tau} d\tau (B+\Delta B)KC] - P \quad (3.1)$$

is negative definite for all $(\Delta A, \Delta B) \in U$. This follows since $x' \Delta V x$ is the Lyapunov difference associated with $V = x'Px$ and (2.3); i.e.,

$$x' \Delta V(\Delta A, \Delta B, P)x = V(x(k+1)) - V(x(k)).$$

System (2.3) is robustly discrete-time stable if this difference is negative for all $x \in \mathbb{R}^n$ and all $(\Delta A, \Delta B) \in U$; e.g., see [7].

A critical step in our development is to express the uncertainty in (3.1) solely in terms of the matrix exponential. Indeed, using the identity (see [8])

$$\exp\left(\begin{bmatrix} (A + \Delta A) & (B + \Delta B) \\ 0 & 0 \end{bmatrix} h\right) = \begin{bmatrix} e^{(A+\Delta A)h} & \int_0^h e^{(A+\Delta A)\tau} d\tau(B+\Delta B) \\ 0 & I \end{bmatrix}, \quad (3.2)$$

(3.1) can be rewritten

$$\Delta V(\Delta A, \Delta B, P) = [I \ 0] \exp\left(\begin{bmatrix} (A+\Delta A) & (B+\Delta B) \\ 0 & 0 \end{bmatrix} h\right) \begin{bmatrix} I \\ KC \end{bmatrix} P [I \ C'K'] \exp\left(\begin{bmatrix} (A+\Delta A) & (B+\Delta B) \\ 0 & 0 \end{bmatrix}' h\right) \begin{bmatrix} I \\ 0 \end{bmatrix} - P \quad (3.3)$$

Our next result provides a parameter-independent upper bound to ΔV . To give this bound, let $\alpha > 0$ be given and define

$$\bar{A}_\alpha \triangleq \left(\begin{bmatrix} A & B \\ 0 & 0 \end{bmatrix} + \frac{\alpha}{2} I \right) \oplus \left(\begin{bmatrix} A & B \\ 0 & 0 \end{bmatrix} + \frac{\alpha}{2} I \right) + \frac{1}{\alpha} \sum_{i=1}^p \begin{bmatrix} A_i & B_i \\ 0 & 0 \end{bmatrix} \otimes \begin{bmatrix} A_i & B_i \\ 0 & 0 \end{bmatrix} \quad (3.4)$$

and

$$\Pi_P \triangleq \begin{bmatrix} I \\ KC \end{bmatrix} P [I \ C'K']. \quad (3.5)$$

Lemma 3.1 (See Appendix for proof): If $P \in R^{n \times n}$ is positive definite and symmetric and $\alpha > 0$, then

$$\Delta V(\Delta A, \Delta B, P) =$$

$$[I \ 0] \exp\left(\begin{bmatrix} (A+\Delta A) & (B+\Delta B) \\ 0 & 0 \end{bmatrix} h\right) \begin{bmatrix} I \\ KC \end{bmatrix} P [I \ C'K'] \exp\left(\begin{bmatrix} (A+\Delta A) & (B+\Delta B) \\ 0 & 0 \end{bmatrix}' h\right) \begin{bmatrix} I \\ 0 \end{bmatrix} - P$$

$$\leq [I \ 0] \text{vec}^{-1} [e^{\bar{A}_\alpha h} \text{vec}(\Pi_P)] \begin{bmatrix} I \\ 0 \end{bmatrix} - P \quad (3.6)$$

for all $(\Delta A, \Delta B) \in U$. VVV

Now, using the right-hand side of (3.6), formally set

$$[I \ 0] \text{vec}^{-1}[e^{\bar{A}_\alpha h} \text{vec}(\Pi_P)] \begin{bmatrix} I \\ 0 \end{bmatrix} - P = -I \quad (3.7)$$

which is equivalent to

$$(M_\alpha - I)\text{vec}(P) = -\text{vec}(I) \quad (3.8)$$

where

$$M_\alpha \triangleq [I \ 0] e^{\bar{A}_\alpha h} \left(\begin{bmatrix} I \\ KC \end{bmatrix} \otimes \begin{bmatrix} I \\ KC \end{bmatrix} \right). \quad (3.9)$$

When does (3.8) have a positive-definite solution? Our next lemma gives a sufficient condition.

Lemma 3.2 (See Appendix for proof): If there exists an $\alpha > 0$ such that

$$\rho(M_\alpha) < 1, \quad (3.10)$$

then (3.8) has a positive-definite solution P. $\forall \forall \forall$

Now, assume (3.10) holds for some $\alpha > 0$. From Lemma 3.2, equation (3.7) has a positive-definite solution P; hence, using (3.6) and (3.7) it follows that

$$\Delta V(\Delta A, \Delta B, P) \leq -I \quad (3.11)$$

for all $(\Delta A, \Delta B) \in U$. We have thus proven the following main result.

Theorem 3.1: If there exists an $\alpha > 0$ such that

$$\rho(M_\alpha) < 1,$$

then the sampled-data system (2.3) is robustly discrete-time stable.

It's important to note that the condition of Theorem 3.1, $\rho(M_\alpha) < 1$, is always satisfied if the "nominal" system is stable and there is no uncertainty; i.e., $\Delta A = 0$ and $\Delta B = 0$. Indeed, taking $\alpha \approx 0$, a straightforward manipulation using (3.2), (3.4), (3.9) and identities (A.1) - (A.4) in the Appendix gives

$$M_\alpha = (e^{Ah} + \int_0^h e^{A\tau} B K C d\tau) \otimes (e^{Ah} + \int_0^h e^{A\tau} B K C d\tau). \quad (3.12)$$

Since the nominal system is stable, then all the eigenvalues of

$$e^{Ah} + \int_0^h e^{A\tau} B K C d\tau$$

lie within the unit disk. This implies, together with (3.12) and the fact that the eigenvalues of a Kronecker product of two matrices are the products of the eigenvalues of these two matrices (see [9]), that all the eigenvalues of M_α lie within the unit disk. Thus, $\rho(M_\alpha) < 1$.

The stability result of Theorem 3.1 is also valid for the case when (2.3) has time-varying uncertain parameters $\sigma_i(t)$. This is a consequence of having established stability via a parameter-independent Lyapunov function. Finally, we remark that a dual result holds when $\Delta B = 0$ and one allows uncertainty in C ; i.e., $C \rightarrow C + \Delta C$.

4. Example

In Soroka and Shaked [13], the robustness of a continuous-time system under "cheap" LQ regulation is studied. The resulting closed-loop system is described by

$$\begin{aligned} \dot{x}(t) &= [A + (B + \Delta B)L_r]x(t) \\ y(t) &= Cx(t) \end{aligned} \quad (4.1)$$

where

$$A = \begin{bmatrix} -1 & 0 \\ 0 & -2 \end{bmatrix}; \quad B = \begin{bmatrix} 1 \\ 1 \end{bmatrix}; \quad C = \begin{bmatrix} -1 \\ 1 \end{bmatrix}';$$

$$\Delta A = 0; \quad \Delta B = \sigma_1 B_1; \quad B_1 = \begin{bmatrix} \bar{\sigma} \\ 0 \end{bmatrix}; \quad \sigma_1^2 \leq 1; \quad \bar{\sigma} > 0$$
(4.2)

and where the LQ regulator gains are given by

$$L_r = [1 + q - \sqrt{5 + 2q} \quad 2\sqrt{5 + 2q} - q - 4]'; \quad q = \sqrt{4 + 1/r}$$
(4.3)

with $r > 0$ being the control weighting in the quadratic performance index

$$\int_0^h [y^2(t) + ru^2(t)] dt.$$

Soroka and Shaked showed that stability robustness decreased as the control became more cheap ($r \rightarrow 0$).

Now, suppose this LQ regulator gain is to be implemented in a computer. What is the robustness of this sampled-data control system? We shall use the sufficient condition in Theorem 3.1 to help answer this question. First, however, we translate the continuous-time LQ gain to one suitable for sampled-data control since the nominal discrete-time system is unstable if we implement gain $K = L_r$ in the computer. Following [pp. 189-191, Astrom and Wittenmark], we take

$$K = L_r [I + (A - BL_r)h/2].$$
(4.4)

For given sampling period h , control weighting r and uncertainty bound $\bar{\sigma}$, we are now in a position to determine if $\rho(M_\alpha) < 1$ for some $\alpha > 0$. For example, for $h = .1$, $r = .05$ and $\bar{\sigma} = .5$, we plot $\rho(M_\alpha)$ versus α in Figure 2 and observe that $\rho(M_\alpha) < 1$ for $\alpha > .3$. Hence, the sampled-data system is stable for $\bar{\sigma} = .5$. Also, for $h = .1$ we determine, for various r , the largest $\bar{\sigma}$ for which $\rho(M_\alpha) < 1$ for some $\alpha > 0$. We compare these results to the actual

discrete-time stability boundary and to Soroka and Shaked's results [13]; see Figure 3. To compute the the actual discrete-time stability boundary, we assume time-invariant uncertainty σ_1 and find the largest $\bar{\sigma}$ for which

$$\rho(e^{Ah} + \int_0^h e^{A\tau} d\tau \begin{bmatrix} 1+\bar{\sigma}_1 \\ 1 \end{bmatrix} K) < 1. \quad (4.5)$$

for all $|\sigma_1| \leq \bar{\sigma}$. Recall that the stability result in Theorem 3.1 is valid for time-varying uncertainties $\sigma_1(t)$ as well; hence, it's natural to expect that the results using $\rho(M_\alpha)$ will be conservative compared to those using (4.5). This was indeed true except for values $1/r \approx 10$. For these values, the stability criterion of Theorem 3.1 predicts ranges of stable σ_1 which are larger than those computed using (4.5). This is clearly incorrect and we contribute these discrepancies to roundoff errors in the computations. Finally, we note that the actual discrete-time stability region is larger than its continuous-time counterpart.

ACKNOWLEDGEMENT

We thank Mr. Mujdat Arabacioglu for his assistance in preparing the example.

APPENDIX: PROOF OF LEMMAS 3.1 AND 3.2

To prove Lemmas 3.1 and 3.2 we first need some identities and observations; see [9] for details.

Identities: If $X \in R^{r \times r}$ and $Y \in R^{s \times s}$, then by definition the Kronecker sum is given by

$$X \oplus Y = X \otimes I + I \otimes Y. \quad (A.1)$$

Next, if the indicated products exist, then

$$\text{vec}(XYZ) = (Z' \otimes X) \text{vec}(Y) \quad (A.2)$$

$$(X \otimes Y)(Z \otimes W) = (XZ) \otimes (YW). \quad (A.3)$$

Finally, if X is square, then

$$e^{X \oplus X} = e^X \otimes e^X. \quad (A.4)$$

Observation 1: Given arbitrary $(\Delta A, \Delta B) \in U$, the solution to the matrix differential equation

$$\dot{Y}(t) = \begin{bmatrix} (A+\Delta A) & (B+\Delta B) \\ 0 & 0 \end{bmatrix} Y(t) + Y(t) \begin{bmatrix} (A+\Delta A) & (B+\Delta B) \\ 0 & 0 \end{bmatrix}', \quad t \geq 0, \quad Y(0) = Y_0 \quad (A.5)$$

is

$$Y(t) = \exp\left(\begin{bmatrix} (A+\Delta A) & (B+\Delta B) \\ 0 & 0 \end{bmatrix} t\right) Y_0 \exp\left(\begin{bmatrix} (A+\Delta A) & (B+\Delta B) \\ 0 & 0 \end{bmatrix}' t\right). \quad (A.6)$$

This is a well-known result.

Observation 2: Let $\alpha > 0$, then the solution to the matrix differential equation

$$\dot{\bar{Y}}(t) = \left(\begin{bmatrix} A & B \\ 0 & 0 \end{bmatrix} + \frac{\alpha}{2} I \right) \bar{Y}(t) + \bar{Y}(t) \left(\begin{bmatrix} A & B \\ 0 & 0 \end{bmatrix}' + \frac{\alpha}{2} I \right) + \frac{1}{\alpha} \sum_{i=1}^p \begin{bmatrix} A_i & B_i \\ 0 & 0 \end{bmatrix} \bar{Y}(t) \begin{bmatrix} A_i & B_i \\ 0 & 0 \end{bmatrix}',$$

$$t \geq 0, \quad \bar{Y}(0) = Y_0$$

(A.7)

is

$$\bar{Y}(t) = \text{vec}^{-1} [e^{\bar{A} t} \text{vec}(\bar{Y}_0)].$$

(A.8)

To show (A.8), apply the "vec" operation to both sides of (A.7) and use identities (A.1) and (A.2) to get

$$\text{vec}(\dot{\bar{Y}}(t)) =$$

$$\left(\left(\begin{bmatrix} A & B \\ 0 & 0 \end{bmatrix} + \frac{\alpha}{2} I \right) \oplus \left(\begin{bmatrix} A & B \\ 0 & 0 \end{bmatrix} + \frac{\alpha}{2} I \right) + \frac{1}{\alpha} \sum_{i=1}^p \left(\begin{bmatrix} A_i & B_i \\ 0 & 0 \end{bmatrix} \otimes \begin{bmatrix} A_i & B_i \\ 0 & 0 \end{bmatrix} \right) \right) \text{vec}(\bar{Y}(t))$$

(A.9)

The solution to (A.9) is

$$\text{vec}(\bar{Y}(t)) = [e^{\bar{A} t} \text{vec}(\bar{Y}_0)]$$

from which (A.8) follows.

Observation 3: If $\alpha > 0$ and if \bar{Y}_0 is positive semi-definite, then the solution in (A.8) satisfies

$$\bar{Y}(t) = \text{vec}^{-1} [e^{\bar{A} t} \text{vec}(\bar{Y}_0)] \geq 0, \quad t \geq 0.$$

(A.9)

To prove this observation, let $S, N \in \mathbb{R}^{r \times r}$ with N positive semi-definite. From (A.2)

$$(S \otimes S) \text{vec}(N) = \text{vec}(SNS')$$

which implies

$$\text{vec}^{-1}[(S \otimes S)\text{vec}(N)] = SNS' \geq 0. \quad (\text{A.10})$$

Using (A.3)

$$(S \otimes S)^i = (S^i \otimes S^i) \quad (\text{A.11})$$

so that

$$\begin{aligned} \text{vec}^{-1}[e^{S \otimes S}\text{vec}(N)] &= \text{vec}^{-1}\left[\sum_{i=0}^{\infty} (i!)^{-1} (S \otimes S)^i \text{vec}(N)\right] \\ &= \text{vec}^{-1}\left[\sum_{i=0}^{\infty} (i!)^{-1} (S^i \otimes S^i) \text{vec}(N)\right] \\ &= \sum_{i=0}^{\infty} \text{vec}^{-1}[(i!)^{-1} (S^i \otimes S^i) \text{vec}(N)] \\ &= \sum_{i=0}^{\infty} (i!)^{-1} S^i N S^{i'} \\ &\geq 0. \end{aligned} \quad (\text{A.12})$$

Furthermore, from (A.4) and (A.10)

$$\begin{aligned} \text{vec}^{-1}[e^{S \oplus S}\text{vec}(N)] &= \text{vec}^{-1}[(e^S \otimes e^S)\text{vec}(N)] \\ &= e^S N e^{S'} \\ &\geq 0. \end{aligned} \quad (\text{A.13})$$

Now, using the exponential product formula (see [10, pg. 97]) we write

$$e^{\bar{A}_\alpha t} = \lim_{j \rightarrow \infty} \left\{ \exp\left[\frac{1}{j}(\tilde{A}_\alpha \oplus \tilde{A}_\alpha)t\right] \prod_{i=1}^p \exp\left[\frac{1}{\alpha_j}(\tilde{A}_i \otimes \tilde{A}_i)t\right] \right\}^j \quad (\text{A.14})$$

where

$$\tilde{A}_\alpha \triangleq \begin{bmatrix} A & B \\ 0 & 0 \end{bmatrix} + \frac{\alpha}{2} I; \quad \tilde{A}_i \triangleq \begin{bmatrix} A_i & B_i \\ 0 & 0 \end{bmatrix}. \quad (\text{A.15})$$

Consequently,

$$\begin{aligned} & \text{vec}^{-1}[e^{\bar{A}_\alpha t} \text{vec}(\bar{Y}_0)] \\ &= \text{vec}^{-1} \left\{ \lim_{j \rightarrow \infty} \left\{ \exp\left[\frac{1}{j}(\tilde{A}_\alpha \oplus \tilde{A}_\alpha)t\right] \prod_{i=1}^p \exp\left[\frac{1}{\alpha_j}(\tilde{A}_i \otimes \tilde{A}_i)t\right] \right\}^j \text{vec}(\bar{Y}_0) \right\} \\ &= \lim_{j \rightarrow \infty} \left\{ \text{vec}^{-1} \left\{ \left\{ \exp\left[\frac{1}{j}(\tilde{A}_\alpha \oplus \tilde{A}_\alpha)t\right] \prod_{i=1}^p \exp\left[\frac{1}{\alpha_j}(\tilde{A}_i \otimes \tilde{A}_i)t\right] \right\}^j \text{vec}(\bar{Y}_0) \right\} \right\}. \end{aligned} \quad (\text{A.16})$$

We'll now show that the expression in the limit brackets in (A.16) is positive semi-definite for all positive integers j . Indeed, this is sufficient to prove that (A.9) holds. For simplicity take $j = p = 1$ and let \tilde{N} satisfy

$$\text{vec}(\tilde{N}) = \exp\left[\frac{1}{\alpha}(\tilde{A}_i \otimes \tilde{A}_i)t\right] \text{vec}(\bar{Y}_0) \quad (\text{A.17})$$

or equivalently

$$\tilde{N} = \text{vec}^{-1} \left\{ \exp\left[\frac{1}{\alpha}(\tilde{A}_i \otimes \tilde{A}_i)t\right] \text{vec}(\bar{Y}_0) \right\}. \quad (\text{A.18})$$

Since \bar{Y}_0 is assumed positive semi-definite, then, from (A.12), \tilde{N} is positive semi-definite. From (A.13), (A.17) and this fact, the expression in the limit brackets of (A.16) satisfies

$$\exp[(\tilde{A}_\alpha \oplus \tilde{A}_\alpha)t] \exp[\frac{1}{\alpha}(\tilde{A}_i \otimes \tilde{A}_i)t] \text{vec}(\bar{Y}_0) = \exp[(\tilde{A}_\alpha \oplus \tilde{A}_\alpha)t] \text{vec}(\tilde{N})$$

$$\geq 0.$$

A similar argument, using (A.12) and (A.13) alternately, shows that the expression in the limit brackets of (A.16) is positive semi-definite for arbitrary positive integers j and p . This completes the proof of Observation 3.

Proof of Lemma 3.1: Assume P positive definite and $\alpha > 0$. Furthermore, consider the matrix differential equations in (A.5) and (A.7) with

$$Y_0 = \bar{Y}_0 = \Pi_P \quad (A.19)$$

where Π_P is given in (3.5). Subtracting (A.5) and (A.7) gives

$$\begin{aligned} \dot{\bar{Y}}(t) - \dot{Y}(t) &= \begin{bmatrix} (A+\Delta A) & (B+\Delta B) \\ 0 & 0 \end{bmatrix} (\bar{Y}(t) - Y(t)) + (\bar{Y}(t) - Y(t)) \begin{bmatrix} (A+\Delta A) & (B+\Delta B) \\ 0 & 0 \end{bmatrix}' + \Psi_\alpha(t) \\ t &\geq 0, \quad \bar{Y}(0) - Y(0) = 0 \end{aligned} \quad (A.20)$$

where

$$\begin{aligned} \Psi_\alpha(t) &= \tilde{A}_\alpha \bar{Y}(t) + \bar{Y}(t) \tilde{A}_\alpha' + \frac{1}{\alpha} \sum_{i=1}^P \tilde{A}_i \bar{Y}(t) \tilde{A}_i' \\ &\quad - \left(\begin{bmatrix} (A+\Delta A) & (B+\Delta B) \\ 0 & 0 \end{bmatrix} \bar{Y}(t) + \bar{Y}(t) \begin{bmatrix} (A+\Delta A) & (B+\Delta B) \\ 0 & 0 \end{bmatrix}' \right) \end{aligned} \quad (A.21)$$

and where \tilde{A}_α and \tilde{A}_i are defined in (A.15).

Claim 1: $\Psi_\alpha(t)$ is positive semi-definite for all $t \geq 0$.

Proof of Claim 1: From (A.15) and (A.21)

$$\begin{aligned}
\Psi_{\alpha}(t) &= \left(\begin{bmatrix} A & B \\ 0 & 0 \end{bmatrix} + \frac{\alpha}{2} I \right) \bar{Y}(t) + \bar{Y}(t) \left(\begin{bmatrix} A & B \\ 0 & 0 \end{bmatrix}' + \frac{\alpha}{2} I \right) \\
&\quad - \left(\begin{bmatrix} (A+\Delta A) & (B+\Delta B) \\ 0 & 0 \end{bmatrix} \bar{Y}(t) + \bar{Y}(t) \begin{bmatrix} (A+\Delta A) & (B+\Delta B) \\ 0 & 0 \end{bmatrix}' \right) + \frac{1}{\alpha} \sum_{i=1}^P \tilde{A}_i \bar{Y}(t) \tilde{A}_i' \\
&= \alpha \bar{Y}(t) - \left(\begin{bmatrix} (A+\Delta A) & (B+\Delta B) \\ 0 & 0 \end{bmatrix} \bar{Y}(t) + \bar{Y}(t) \begin{bmatrix} (A+\Delta A) & (B+\Delta B) \\ 0 & 0 \end{bmatrix}' \right) + \frac{1}{\alpha} \sum_{i=1}^P \tilde{A}_i \bar{Y}(t) \tilde{A}_i' \\
&\geq \alpha \sum_{i=1}^P \sigma_i^2 \bar{Y}(t) - \sum_{i=1}^P \sigma_i (\tilde{A}_i \bar{Y}(t) + \bar{Y}(t) \tilde{A}_i') + \frac{1}{\alpha} \sum_{i=1}^P \tilde{A}_i \bar{Y}(t) \tilde{A}_i' \\
&= \sum_{i=1}^P \left[\sqrt{\alpha} \sigma_i I - \frac{1}{\sqrt{\alpha}} \tilde{A}_i \right] \bar{Y}(t) \left[\sqrt{\alpha} \sigma_i I - \frac{1}{\sqrt{\alpha}} \tilde{A}_i \right]'.
\end{aligned} \tag{A.22}$$

From Observation 3, $\bar{Y}(t)$ is positive semi-definite for all $t \geq 0$. It thus follows from (A.22) that $\Psi_{\alpha}(t)$ is likewise positive semi-definite for all $t \geq 0$. This proves the claim.

Now, since $\bar{Y}(0) - Y(0) = 0$, the solution to (A.20) is

$$\bar{Y}(t) - Y(t) = \int_0^t \exp \left[\begin{bmatrix} (A+\Delta A) & (B+\Delta B) \\ 0 & 0 \end{bmatrix} (t-s) \right] \Psi_{\alpha}(s) \exp \left[\begin{bmatrix} (A+\Delta A) & (B+\Delta B) \\ 0 & 0 \end{bmatrix}' (t-s) \right] ds. \tag{A.23}$$

From (A.23) and Claim 1 it follows that

$$\bar{Y}(t) - Y(t) \geq 0, \quad t \geq 0. \tag{A.24}$$

Combining (A.6) and (A.8) with (A.24) gives

$$\begin{aligned}
0 &\leq \bar{Y}(h) - Y(h) \\
&= \text{vec}^{-1} \left[e^{\tilde{A} h} \text{vec}(\Pi_P) \right] - \exp \left(\begin{bmatrix} (A+\Delta A) & (B+\Delta B) \\ 0 & 0 \end{bmatrix} h \right) \Pi_P \exp \left(\begin{bmatrix} (A+\Delta A) & (B+\Delta B) \\ 0 & 0 \end{bmatrix}' h \right)
\end{aligned}$$

which implies that

$$[I \ 0] \exp \left(\begin{bmatrix} (A+\Delta A) & (B+\Delta B) \\ 0 & 0 \end{bmatrix} h \right) \Pi_P \exp \left(\begin{bmatrix} (A+\Delta A) & (B+\Delta B) \\ 0 & 0 \end{bmatrix}' h \right) \begin{bmatrix} I \\ 0 \end{bmatrix} = P$$

$$\leq [I \ 0] \text{vec}^{-1} [e^{\bar{A}_\alpha h} \text{vec}(\Pi_p)] \begin{bmatrix} I \\ 0 \end{bmatrix} - P$$

which is the desired result. The proof of Lemma 3.1 is complete. $\forall \forall \forall$

Proof of Lemma 3.2: Assume $\alpha > 0$ such that (3.10) holds. We must show that (3.8) has a positive-definite solution P . Indeed, since $\rho(M_\alpha) < 1$, then $(M_\alpha - I)$ is invertible and (3.8) has a unique solution

$$\text{vec}(P) = (I - M_\alpha)^{-1} \text{vec}(I)$$

or

$$P = \text{vec}^{-1}[(I - M_\alpha)^{-1} \text{vec}(I)].$$

Now, with $M_\alpha^0 \triangleq I$, it follows from [11, Theorem 6.7.1] that $(I - M_\alpha)^{-1} = \sum_{i=0}^{\infty} M_\alpha^i$. Consequently,

$$P = \text{vec}^{-1} \left[\sum_{i=0}^{\infty} M_\alpha^i \text{vec}(I) \right] = \sum_{i=0}^{\infty} \text{vec}^{-1} [M_\alpha^i \text{vec}(I)]. \quad (\text{A.25})$$

Claim 2: If i is a positive integer, then

$$\text{vec}^{-1} [M_\alpha^i \text{vec}(I)] \geq 0. \quad (\text{A.26})$$

Proof of Claim 2: The proof proceeds by induction. From the definition of M_α in (3.9) and identity (A.2), we have for $i = 1$

$$\begin{aligned} \text{vec}^{-1} [M_\alpha \text{vec}(I)] &= \text{vec}^{-1} \{ [I \ 0] e^{\bar{A}_\alpha h} \left(\begin{bmatrix} I \\ KC \end{bmatrix} \otimes \begin{bmatrix} I \\ KC \end{bmatrix} \right) \text{vec}(I) \} \\ &= \text{vec}^{-1} \{ [I \ 0] e^{\bar{A}_\alpha h} \text{vec} \left(\begin{bmatrix} I \\ KC \end{bmatrix} [I \ C'K'] \right) \} \\ &= [I \ 0] \text{vec}^{-1} \{ e^{\bar{A}_\alpha h} \text{vec} \left(\begin{bmatrix} I \\ KC \end{bmatrix} [I \ C'K'] \right) \} \begin{bmatrix} I \\ 0 \end{bmatrix}. \end{aligned} \quad (\text{A.27})$$

It follows from (A.9), (A.27) and Observation 3 with

$$\bar{Y}(0) = \begin{bmatrix} I \\ KC \end{bmatrix} [I \ C'K'], \quad t = h$$

that $\text{vec}^{-1}[M_{\alpha} \text{vec}(I)] \geq 0$. Now, for the induction step, assume $\text{vec}^{-1}[M_{\alpha}^i \text{vec}(I)] \geq 0$. Then

$$\begin{aligned} \text{vec}^{-1}[M_{\alpha}^{i+1} \text{vec}(I)] &= \text{vec}^{-1}[M_{\alpha} M_{\alpha}^i \text{vec}(I)] \\ &= \text{vec}^{-1}\{M_{\alpha} \text{vec}[\text{vec}^{-1}(M_{\alpha}^i \text{vec}(I))]\}. \end{aligned} \quad (\text{A.28})$$

From (A.28) and Observation 3 with

$$\bar{Y}(0) = \text{vec}^{-1}[M_{\alpha}^i \text{vec}(I)], \quad t = h,$$

it follows that $\text{vec}^{-1}[M_{\alpha}^{i+1} \text{vec}(I)] \geq 0$. This proves the lemma.

vvv

REFERENCES

- [1] J. Manela, "Deterministic Control of Uncertain Linear Discrete and Sampled-Data Systems," Ph.D. Dissertation, Univ. Calif., Berkeley, 1985.
- [2] M. Corless, "Stabilization of Uncertain Linear Systems," Proceedings of the IFAC Workshop on Model Error Concepts and Compensation, Boston, 1985.
- [3] C.V. Hollot and M. Arabacioglu, "lth-step Lyapunov Min-Max Controllers: Stabilizing Discrete-Time Systems Under Real Parameter Variations," Proceedings of the American Control Conference, Minneapolis, pp. 496-501, 1987.
- [4] M.E. Magana and S. Zak, "Robust State Feedback Stabilization of Discrete-Time Uncertain Dynamical Systems," IEEE Transactions on Automatic Control, Vol. AC-33, no. 9, pp. 887-891, 1988.
- [5] P.M. Thompson, R.L. Dailey and J.C. Doyle, "New Conic Sectors for Sampled-Data Feedback Systems," Systems and Control Letters, Vol. 7, pp. 395-404, 1986.
- [6] A.R. Tiedemann and W.L. De Koning, "The Equivalent Discrete-Time Optimal Control Problem for Continuous-Time Systems with Stochastic Parameters," International Journal of Control, Vol. 40, no. 3, pp. 449-466, 1984.
- [7] K.J. Astrom and B. Wittenmark, Computer Controlled Systems, Theory and Design, Prentice-Hall, Englewood Cliffs, NJ, 1984.
- [8] C.F. Van Loan, "Computing Integrals Involving the Matrix Exponential," IEEE Transactions on Automatic Control, Vol. AC-20, no. 3, p. 395-404, 1978.
- [9] J.W. Brewer, "Kronecker Products and Matrix Calculus in System Theory," IEEE Transactions on Circuits and Systems, Vol. CAS-25, no. 9, pp. 772-781, 1978.
- [10] V.S. Varadarajan, Lie Groups, Lie Algebras, and Their Representations, Springer-Verlag, New York, 1984.
- [11] A.W. Naylor and G.R. Sell, Linear Operator Theory in Engineering and Science, Holt, Rinehart and Winston, Inc., New York, 1971.
- [12] D.S. Bernstein, "Robust Static and Dynamic Output-Feedback Stabilization: Deterministic and Stochastic Perspectives," IEEE Transactions on Automatic Control, Vol. AC-32, no. 12, pp. 1076-1084, 1987.
- [13] E. Soroka and U. Shaked, "On the Robustness of LQ Regulators," IEEE Transactions on Automatic Control, Vol. AC-29, no. 7, pp. 664-665, 1984.

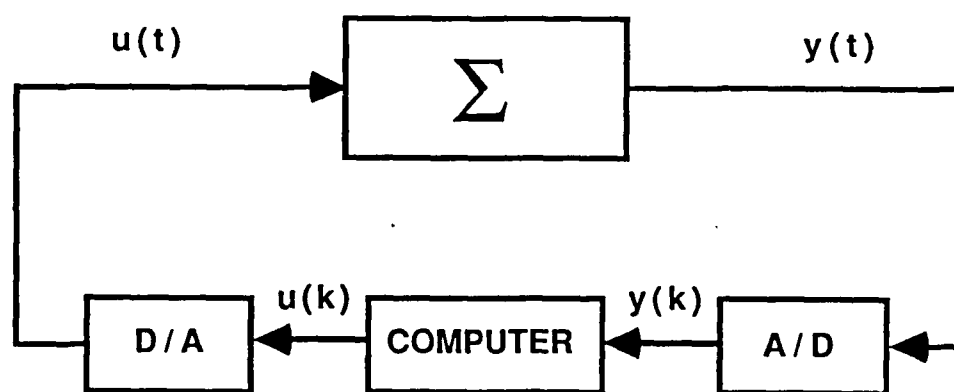


Figure 1 : Sampled-Data Control System

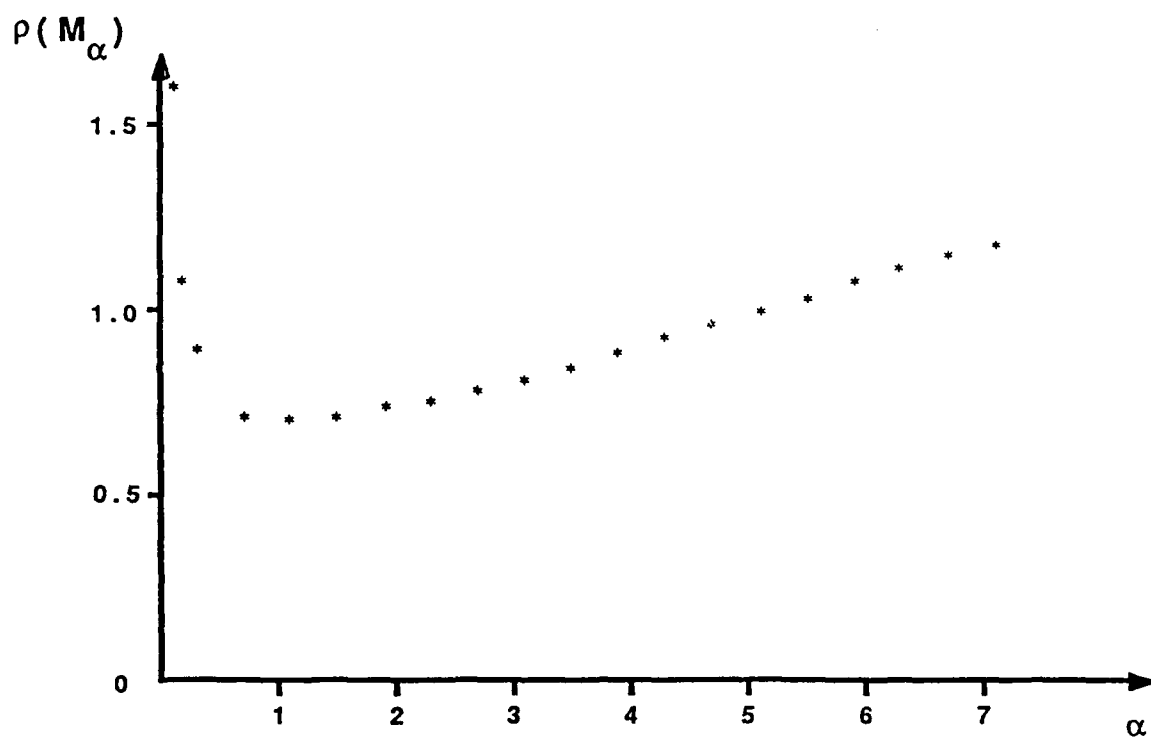


Figure 2 : Plot of $\rho(M_\alpha)$ versus α for $h=0.1$,
 $r=0.05$, $\bar{\sigma}=0.5$

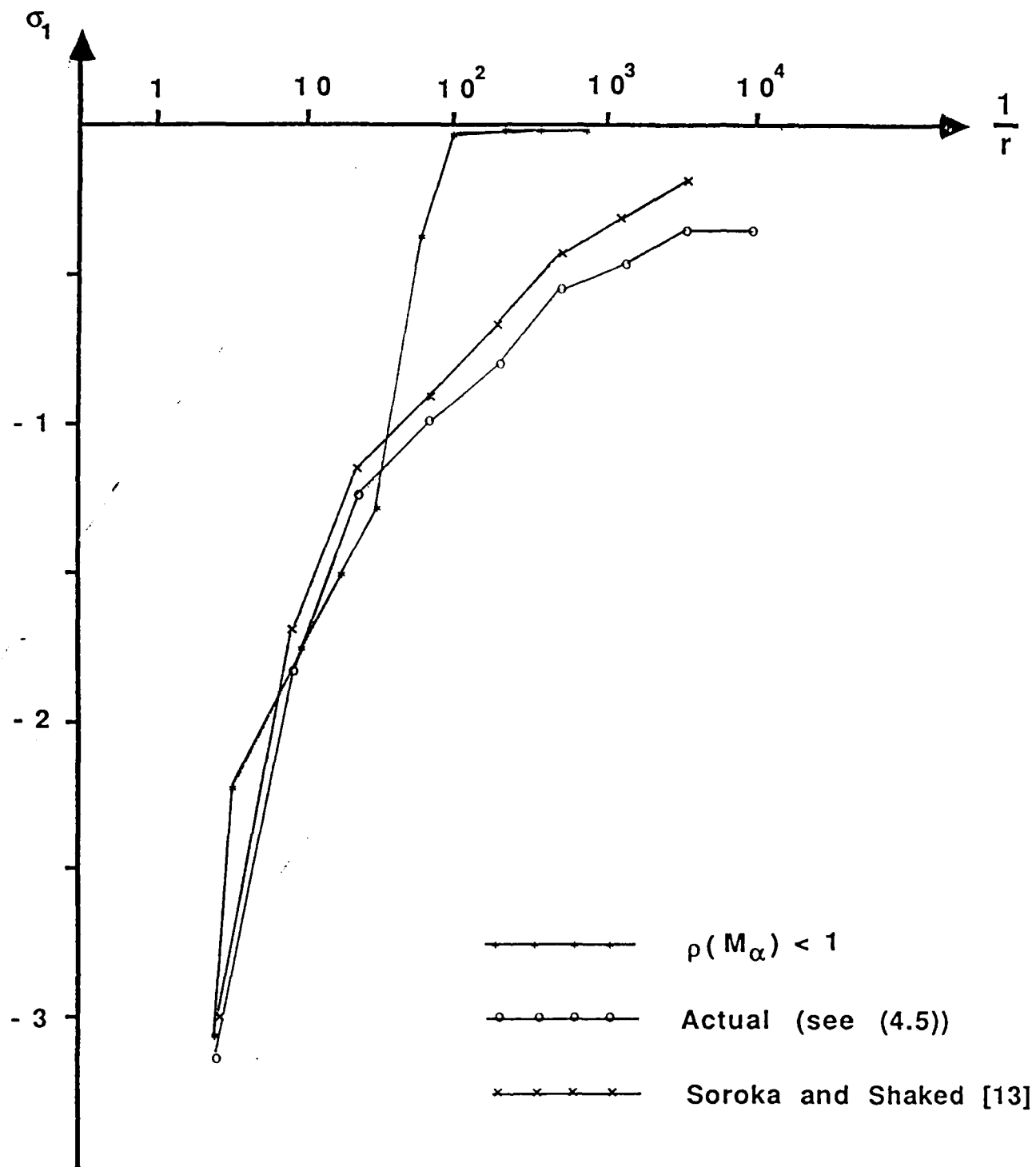


Figure 3 : Comparison of several methods for
determining the stability boundary